



HAL
open science

Structures et aléa en finance, une approche par la complexité algorithmique de l'information

Lin Ma

► **To cite this version:**

Lin Ma. Structures et aléa en finance, une approche par la complexité algorithmique de l'information. Intelligence artificielle [cs.AI]. Université des Sciences et Technologie de Lille - Lille I, 2010. Français. NNT: . tel-00839386

HAL Id: tel-00839386

<https://theses.hal.science/tel-00839386v1>

Submitted on 27 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Structures et aléa en finance, une approche par la complexité algorithmique de l'information

THÈSE

présentée et soutenue publiquement le 23/11/2010

pour l'obtention du

Doctorat de l'Université Lille 1 - Sciences et Technologies
(Spécialité Sciences de Gestion)

par

Lin MA

Composition du jury

<i>Rapporteurs :</i>	Franck Moraux	Professeur des Universités IAE - Université Rennes 1
	Gilles Richard	Professeur des Universités Université Toulouse 3 - Paul Sabatier
<i>Examineurs :</i>	Jérôme Maati	Professeur des Universités Université Lille 1 - Sciences et Technologies
	Philippe Mathieu	Professeur des Universités Université Lille 1 - Sciences et Technologies
	Hélène Rainelli	Professeur des Universités IAE - Université Paris 1 - Panthéon Sorbonne
<i>Directeurs :</i>	Olivier Brandouy	Professeur des Universités IAE - Université Paris 1 - Panthéon Sorbonne
	Jean-Paul Delahaye	Professeur des Universités Université Lille 1 - Sciences et Technologies

UNIVERSITÉ LILLE 1 - SCIENCES ET TECHNOLOGIES

Lille Economie et Management — UMR 8179

104, avenue du peuple belge — 59043 LILLE Cedex

Tél. : +33 (0)3.20.12.34.44

REMERCIEMENTS

Je tiens tout d'abord à remercier mes directeurs de thèse, Professeur Olivier Brandouy et Professeur Jean-Paul Delahaye, pour leur encadrement rigoureux et efficace. La nature interdisciplinaire de cette thèse implique beaucoup d'investissements personnels de leur part, c'est leurs qualités en tant que directeur de thèse qui m'ont permis d'avancer rapidement dans la bonne direction. Leur oeil critique m'a été très précieux pour structurer le travail et pour améliorer la qualité des différentes sections. Que ces lignes leur assurent de ma très sincère reconnaissance.

Je remercie également les doctorants, maîtres de conférences et professeurs, que j'ai eu la chance de côtoyer au sein du laboratoire Lille Economie et Management (LEM). Leurs conseils ont contribué à l'aboutissement de ce travail autant sur le fond que sur la forme. Je remercie en particulier Fabrice pour la relecture du document final.

Un grand merci à tous mes amis de Lille, Géraldine, Sophie, Candy, Antoine, Aurély, Emilien, et Zhen dont les encouragements et amitiés ont enrichi ce parcours de thèse autant sur le plan social qu'intellectuel.

Une pensée reconnaissante et affectueuse pour Cédric et mes familles chinoise et française dont l'affection et le soutien indéfectibles m'ont permis de surmonter les étapes les plus difficiles de ce travail.

“*Last but not least*”, je remercie vivement Madame et Messieurs les membres du jury d'avoir accepté d'évaluer ce travail dont les critiques précieuses contribueront certainement à son amélioration.

TABLE DES MATIÈRES

<i>Introduction générale</i>	1
<i>1. Hasard, structure et prévisibilité : un point sur les théories en finance</i>	9
Introduction	10
1.1 Imprévisibilité des rentabilités excessives	12
1.1.1 Rentabilité espérée, anticipée, exigée et d'équilibre, les moteurs du mécanisme de marché.	13
1.1.2 De l'efficience informationnelle à la finance conventionnaliste : peut-on prévoir les rentabilités excessives ?	15
1.1.3 L'efficience informationnelle, comment la tester ?	27
1.2 Absence de consensus sur le modèle d'équilibre, une source de confusion	28
1.2.1 Du CAPM au modèle multi-facteurs, comment calculer la rentabilité exigée ?	28
1.2.2 Limites théoriques des fonds d'investissement : dans quelle mesure les marchés financiers sont-ils imbattables ?	38
1.3 Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives	43
1.3.1 Effet calendrier	44

Table des matières

1.3.2	Effet “momentum”	46
1.3.3	Excès de la volatilité et retour de moyenne	47
1.3.4	Prime de risque négative	49
1.3.5	Sur-réaction et sous-réaction événementielles	50
	Conclusion	53
2.	<i>Régularités empiriques sur les marchés boursiers</i>	55
	Introduction	56
2.1	Propriétés statistiques des séries de rentabilités	57
2.1.1	Caractérisation de la moyenne	58
2.1.2	Propriétés des moments d’ordre supérieur à deux	74
2.1.3	Propriétés des rentabilités financières : un récapitulatif	86
2.2	De l’approche économétrique à l’impossibilité de “ <i>market timing</i> ”	87
2.2.1	Les modèles de la famille ARMA	87
2.2.2	ARCH, GARCH, et FIGARCH	88
2.2.3	Modèles multi-états	89
2.2.4	Retour sur l’impossibilité de “ <i>market timing</i> ”	92
	Conclusion	95
3.	<i>Complexité de Kolmogorov : une application de la théorie algorithmique de l’information aux séries de rentabilités</i>	98
	Introduction	99

Table des matières

3.1	Le hasard et la science, un couple d'adversaires?	100
3.2	Une approche probabiliste	102
3.3	Une approche algorithmique	104
3.3.1	Hasard et tests statistiques	104
3.3.2	Hasard et stratégies de pari	108
3.3.3	Hasard et programmes de compression	110
3.4	Les algorithmes de compression : des détecteurs de régularités . .	114
3.4.1	Codages d'entropie	115
3.4.2	Codages par contexte : algorithmes de compression dotés d'une intelligence artificielle.	124
3.5	Détection des régularités financières avec les algorithmes de com- pression	130
3.5.1	Processus de discrétisation : un prétraitement indispensable aux rentabilités financières	133
3.5.2	Effacement des régularités à l'aide des transformations al- gébriques	135
3.5.3	Algorithmes de compression vs. tests statistiques	139
3.5.4	Limites pratiques de la méthode algorithmique : le cas des décimales de π	143
3.5.5	Une discussion méthodologique	145
3.6	Une première application sur les données réelles : le cas de l'indice Dow Jones	150
3.6.1	Détection des faits stylisés	150
3.6.2	Degré d'aléa sans faits stylisés	154

3.7	Application de la méthode aux marchés internationaux	156
3.7.1	Présentation des données	156
3.7.2	Tests de compression et résultats obtenus	158
3.7.3	Traitement des rentabilités zéros	159
3.7.4	Méthodes et analyses	161
3.7.5	Résultats et discussions	172
	Conclusion	175
4.	<i>Une définition computationnelle des séries de cours imbattables</i>	178
	Introduction	179
4.1	Modèle “tout acheter et tout vendre”	182
4.1.1	Différence de gains en termes de rentabilités	186
4.1.2	Propriété partagée par les cours battables	188
4.1.3	Condition d’existence des cours imbattables	189
4.1.4	L’aspect aléatoire d’une suite batable	190
4.1.5	Différence théorique entre “faiblement battre” et “fortement battre”	194
4.1.6	Liens entre des variables <i>i.i.d.</i> et les cours imbattables	196
4.2	Politique du blocage à 1 centime d’euro	201
4.3	Politique de blocage et nature imbattable des cours infinis	206
4.4	Généralisation du modèle “tout acheter et tout vendre”	210
4.5	Stratégies simples vs. stratégies généralisées	212

Table des matières

4.6	La vitesse de convergence des stratégies d'investissement	213
4.6.1	La performance à long terme et la vitesse de convergence, deux aspects complémentaires d'une stratégie d'investisse- ment	216
4.6.2	Cas où la volatilité n'augmente pas dans le temps	220
	Conclusion	222
	<i>Conclusion générale</i>	224
	<i>Annexe</i>	228
	<i>A. Efficience des marchés financiers dans un cadre d'information imparfaite et d'agents hétérogènes</i>	229
	<i>B. "Cross section tests" sur la validité du CAPM</i>	231
	<i>C. Description du test GRS</i>	233
	<i>D. Présentation des trois modèles comportementaux</i>	235
	<i>E. Origine théorique du test de volatilité et du phénomène de retour à la moyenne</i>	237
	<i>F. Statistique BDS</i>	239
	<i>G. Originalité des suites compressibles</i>	240
	<i>H. Compression de la suite de Champernowne</i>	241

Table des matières

<i>I. Analyses statistiques de la distribution $N(0, 1)$</i>	242
<i>J. Différence entre un jeu équitable et un processus indépendant</i>	243
<i>K. Quelques notions élémentaires en théorie de la calculabilité</i>	244
K.1 Thèse de Turing-Church et machine de Turing universelle	246
K.2 Ensemble récursif et ensemble récursivement énumérable	247
K.3 Fonction calculable	249
K.4 Ensemble de mesure nulle	249
<i>L. Quelques techniques complémentaires</i>	251
<i>M. Résultats de compression obtenus sur les données internationales</i>	255
<i>N. Analyse en composantes principales (ACP)</i>	259

INTRODUCTION GÉNÉRALE

Fin 1999, la capitalisation boursière des marchés financiers a dépassé le PIB mondial pour la première fois de l'histoire. Bien que sa fonction économique soit fréquemment remise en question, le pouvoir de la finance ne cesse d'augmenter au regard de l'économie réelle. La financiarisation des économies, après la crise des “*subprimes*” en 2008, a fini par entraîner l'économie mondiale dans la plus sévère récession depuis 1929.

Serait-il possible de prédire un tel scénario catastrophe et de sortir du marché juste avant la chute boursière ? Telle est la question que se pose fatalement les investisseurs après chaque krach inopiné.

Si la possibilité de prédire des crises financières n'est en rien simple, ni même attestée rigoureusement, des investisseurs renommés, tels Warren Buffet, George Soros ou encore James Simons, semblent bien démontrer une certaine aptitude à anticiper les mouvements du marché. Par exemple, le “*Medallion fund*” dirigé par James Simons, professeur en mathématiques, a enregistré un rendement de 80% en 2008, alors que l'indice Dow-Jones baissait de 30,5% pendant la même période. Ces icônes de l'industrie financière, jouissent-elles d'une renommée exagérée ? Comment peut-on interpréter leur performance d'un point de vue théorique ? Répondre à ces questions nécessite une étude complète sur la prévisibilité des dynamiques boursières.

Or, fixer les limites théoriques aux performances des fonds d'investissement, comprendre dans quelle mesure on peut compter sur un gérant de portefeuille, limiterait peut-être la survenue de scandales comme celui provoqué par Bernard Madoff¹ (une escroquerie financière qui doit sa réussite, au moins en partie, à des promesses de rendements fantastiques, pour plus de détails sur le mécanisme mathématique des escroqueries du type Ponzi voir Delahaye (2009a)).

En ce sens, comprendre les mécanismes sous-jacents aux évolutions boursières, notamment l'origine de leur imprévisibilité, semble plus que jamais nécessaire.

1. Bernard Madoff, fondateur de l'entreprise financière “*Bernard L. Madoff Investment Securities LLC*”, l'auteur d'un détournement de fonds de 65 milliards dollars dont les rendements promis devaient atteindre 47 milliards.

C'est pourtant une entreprise déjà ancienne :

La théorie de l'efficience (Fama 1970), définie par l'intégration instantanée des informations pertinentes dans les prix successifs, offre un cadre général aux études des dynamiques boursières. Selon cette théorie, les rentabilités excessives, définies par la différence entre les rentabilités réalisées et d'équilibre, sont imprévisibles et statistiquement inexistantes dans le sens où leur espérance conditionnelle est toujours égale à zéro.

Or, après chaque krach boursier, de violentes critiques accablent cette théorie. A priori, un marché qui fonctionne "bien" devrait évaluer les titres échangés à leur juste valeur, les scénarios catastrophes comme celui de la crise des "*subprimes*" constitueraient pour d'aucuns une preuve empirique d'un dysfonctionnement général et d'une inadéquation de la théorie.

Pour expliquer la reproduction des crises financières, des théories concurrentes, telles que la finance comportementale (Tversky et Kahneman 1974) et l'approche conventionnaliste (Orléan 1999), s'appuient sur la nature irréaliste de la rationalité classique. En effet, selon ces dernières théories, des impulsions psychologiques ou des conventions collectives peuvent pousser le prix d'un actif donné loin de sa valeur fondamentale. Les bulles ou krachs financiers seraient alors des conséquences de ces erreurs d'évaluation.

Or, si leurs différences conceptuelles s'avèrent manifestes, les trois approches citées ci-dessus restent difficiles à départager sur le plan empirique. Aussitôt qu'une étude tente de rejeter l'efficience informationnelle, de nombreux travaux contesteront sa conclusion en adoptant une autre interprétation des faits constatés (Lo et MacKinlay 2002, Malkiel 2003). Ce problème, sans doute dû à la difficulté associée au calcul de la rentabilité d'équilibre (dit "problème de l'hypothèse jointe"), constitue une source d'ambiguïté dans la plupart des tests d'efficience (Fama 1976). La partie théorique de cette thèse est donc destinée à recenser les points de convergence et de divergence entre ces trois théories concurrentes, notamment en ce qui concerne leur implication pour les dynamiques financières, et surtout au regard de leur hypothétique prévisibilité.

En raison de l'évolution incertaine des cours boursiers, les séries financières sont souvent étudiées dans un cadre probabiliste, et les tests statistiques sont les principaux outils disponibles pour caractériser les données historiques.

Or, en mathématiques, la notion d'aléa est aussi étudiée par la théorie de la calculabilité, une branche de l'informatique fondamentale qui examine la possibilité de résoudre un problème donné de façon automatique (Church 1936, Turing 1936). Un des concepts phares de cette dernière théorie, la complexité de Kolmogorov

(Kolmogorov 1965), est reconnu comme l'indicateur d'aléa le plus général en informatique². Pour autant, ce concept puissant est à notre connaissance, peu utilisé en finance. Un des objectifs de cette thèse est d'en établir la pertinence au regard de la discipline financière.

La complexité de Kolmogorov d'une suite s est définie par "la longueur du plus petit programme qui l'engendre". Une méthode pour évaluer cette longueur du plus petit programme est l'utilisation des algorithmes de compression qui représentent s avec le moins de mémoire possible. Étant donné qu'une série est d'autant plus imprévisible que sa complexité de Kolmogorov est grande, son degré d'aléa pourra être estimé par le meilleur résultat de compression obtenu sur elle.

En suivant cette logique, nous estimons, dans cette thèse, le degré d'aléa de diverses séries financières en appliquant des algorithmes de compression. L'introduction de cette méthode pourra approfondir notre compréhension des dynamiques boursières, car certaines expériences montrent qu'il existe des régularités détectables par les algorithmes de compression qui ne le sont pas par les tests statistiques traditionnels, ce qui suggère l'existence de "régularités profondes" (probablement très difficiles à expliquer) dans les séries de cours.

D'un point de vue managérial, comprendre dans quelle mesure les rentabilités sont imprévisibles revient à dessiner les limites théoriques des performances pour les fonds d'investissement. Que veut dire par exemple "battre le marché" ? Pourquoi est-ce impossible pour certaines théories financières ? Nous montrons dans cette thèse que battre l'indice du marché en pratiquant une stratégie de "*stock picking*" est plus raisonnable que de souhaiter battre la "*buy and hold*" en pratiquant le "*market timing*", et ce, aussi bien sur le plan empirique que théorique. Ce résultat pourrait avoir d'importantes implications sur le management des fonds d'investissement, notamment sur l'évaluation de leurs performances.

Mesurer la quantité d'aléa cachée derrière chaque série de rentabilités permet aussi de comparer différentes places boursières et de saisir la relation entre leur mode de fonctionnement et la prévisibilité des rentabilités futures. Par exemple, en supposant qu'un marché est d'autant plus efficient que son niveau d'imprévisibilité est élevé, on pourrait évaluer l'impact d'une mesure microstructurelle en étudiant l'évolution du degré d'aléa avant et après son introduction.

Cette thèse contient quatre chapitres dont chacun constitue un essai qui traite un aspect différent des régularités financières. Pour illustrer l'enchaînement des idées générales qui parcourent ce travail, l'organisation de la thèse est schématisée dans la figure 0.1.

². La définition de la complexité de Kolmogorov est présentée en détail dans la section 3.3.3 page 110.

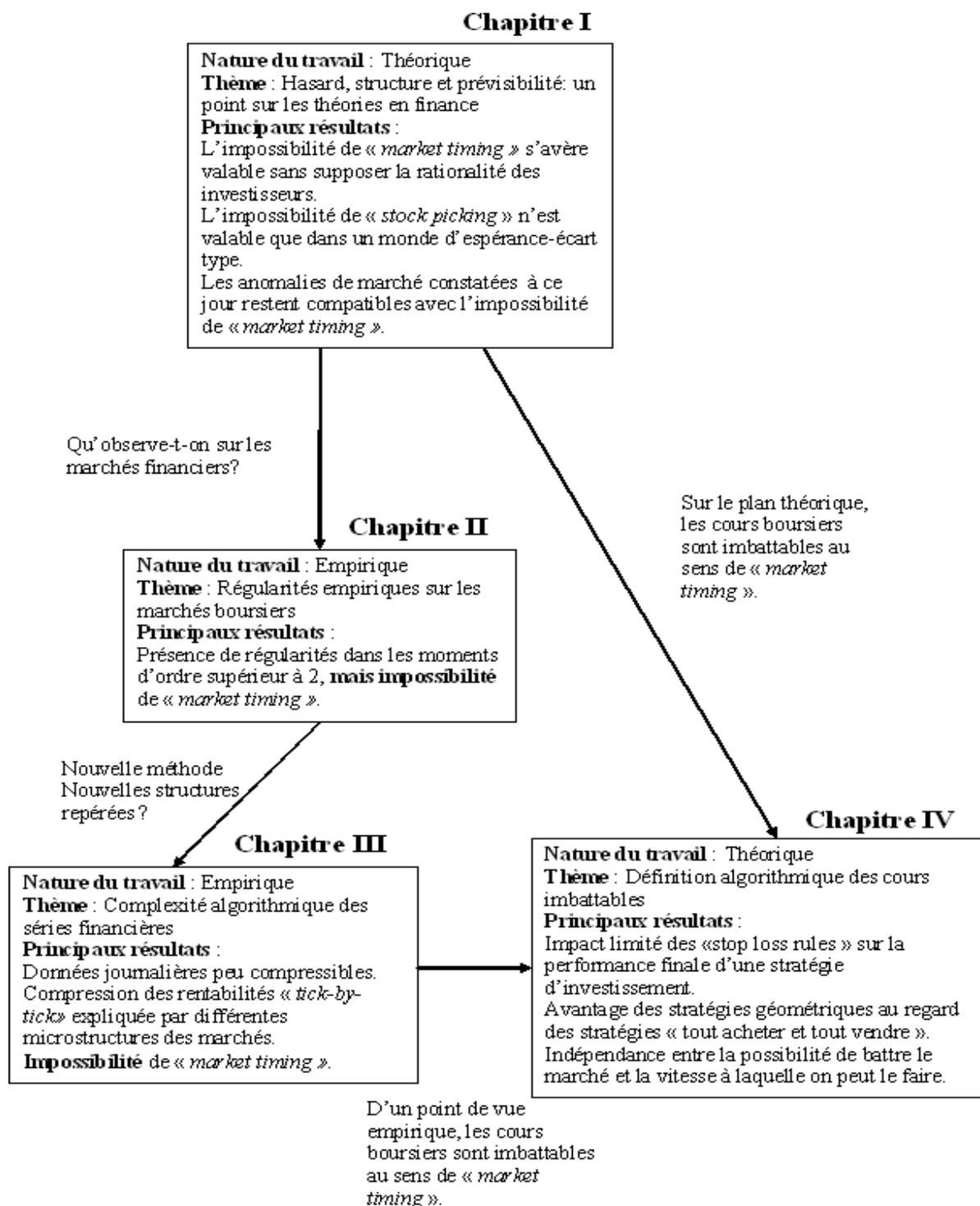


Fig. 0.1: Organisation générale de la thèse

Hasard, structure et prévisibilité : un point sur les théories en finance

L'absence des rentabilités excessives, souvent considérée comme une conséquence directe de la théorie de l'efficience, est en réalité compatible avec les trois théories étudiées dans cette thèse, à savoir, l'efficience informationnelle, la finance comportementale et l'approche conventionnaliste. Sur un marché efficient, les rentabilités excessives, dirigées par les "surprises" en informations fondamentales, sont toujours d'espérance nulle. Contrairement à ce que soutiennent certains auteurs, il ne semble pas que la prise en compte des émotions, de l'humeur ou de l'opinion des investisseurs permet une meilleure prévisibilité. La rationalité des investisseurs, qui se traduit, selon Fama (1976), par l'usage exclusif des informations fondamentales, devient en effet le seul point distinct des trois approches analysées.

Bien que le calcul exact des rentabilités d'équilibre restait controversé, les investisseurs n'ont aucune raison de détenir des titres présentant des rentabilités négatives. Cette positivité des rentabilités espérées implique l'impossibilité de surperformer les stratégies de type "*buy and hold*" avec des stratégies fondées sur le "*market timing*". En revanche, l'inefficacité des stratégies de "*stock picking*" est liée au CAPM, un modèle d'équilibre spécifique valable uniquement dans un monde d'espérance-écart type. En insistant sur cette différence entre l'efficience et le CAPM, les principales anomalies de marché sont étudiées pour défendre leur compatibilité avec l'inefficacité de "*market timing*".

Régularités empiriques sur les marchés boursiers Après avoir mis en évidence son fondement théorique, l'absence des rentabilités excessives est examinée sur le plan empirique. À cet effet, dans le deuxième chapitre, on étudie les propriétés statistiques des séries financières et les modèles économétriques qui peuvent les caractériser.

Bien que les rentabilités successives ne soient pas parfaitement indépendantes, les structures qui les lient sont identifiées uniquement dans leurs moments d'ordre supérieur à 2. Ces régularités, nommées "faits stylisés" (i.e., queues épaisses, bouffées de volatilité, multi-échelle), sont donc présentées dans l'objectif de soutenir l'impossibilité de battre la "*buy and hold*".

L'exploitation des modèles économétriques ne semble pas permettre non plus engendrer des rentabilités excessives, car aucun de ces modèles ne peut prédire le sens des variations futures avec un taux de réussite supérieur à 60%, ce qui selon Grabbe (1996) constitue le seuil minimal à atteindre pour construire une stratégie de "*market timing*" rentable. Ces résultats empiriques restent donc en accord avec les analyses du premier chapitre.

Or, un test statistique est toujours conçu pour un type de régularité bien spécifique; il serait donc tout à fait possible que certaines régularités soient ignorées par les tests traditionnels. En introduisant de nouvelles approches empiriques, on pourrait découvrir d'autres structures dans les séries financières, ou a tout le moins, en déceler la présence sans nécessairement les caractériser.

Complexité algorithmique des séries financières Dans le troisième chapitre, on propose de détecter des régularités financières avec des algorithmes de compression. Plus précisément, une approche générale est introduite pour estimer la complexité de Kolmogorov des séries de rentabilités. Après un processus de “discrétisation” qui transforme les rentabilités successives en suites de nombres entiers, les algorithmes de compression sont utilisés pour détecter des structures régulières parfois invisibles par des tests statistiques. Selon notre principe d’estimation, une suite est d’autant plus prévisible que son taux de compression est important.

Cette approche algorithmique est d’abord illustrée avec des séries de cours simulées, puis appliquée aux suites de rentabilités réelles.

Dans un premier temps, après suppression des régularités inexploitable d’un point de vue de “*market timing*”, les variations journalières de l’indice Dow Jones industriel s’avèrent peu compressibles, donc fortement aléatoires. Ensuite, en appliquant cette même approche aux données *tick-by-tick* observées sur les principaux marchés internationaux (i.e. NYSE, Nasdaq, Euronext-Paris), on constate un degré d’aléa plus élevé sur le marché français que sur les deux bourses américaines. Nous expliquons ce résultat par le fait que l’auto-corrélation des volatilités inter-journalières paraît plus élevée aux Etats-Unis qu’en France. Peu exploitable par des stratégies de “*market timing*”, cette régularité reste compatible avec l’impossibilité de battre une stratégie de type “*buy and hold*”.

Bien que les régularités identifiées dans ce chapitre soient conformes aux théories classiques en finance, l’approche algorithmique ici introduite ouvre de nouvelles perspectives pour étudier les séries des rentabilités. En parfait parallèle avec les tests statistiques, cette approche élargit la gamme des régularités détectables en finance, et en développant de nouveaux outils de compression, elle pourrait révéler des structures novatrices qui modifieraient notre compréhension des marchés financiers.

Définition computationnelle des séries de cours imbattables Si “battre le portefeuille de marché” signifie obtenir un ratio rentabilité/risque plus élevé que celui du marché, l’impossibilité de battre la “*buy and hold*”, seule implication directe de l’efficience informationnelle, n’a jamais fait à notre connaissance l’objet d’une définition mathématique. Dans le quatrième essai de cette thèse, on définit les cours imbattables dans ce sens spécifique en s’inspirant des suites aléatoires (composées de “0” et de “1”) au sens de Schnorr (1971), où les stratégies exécutables sont modélisées par des fonctions calculables³. Une série de cours est dite “battue” par une stratégie effective, $d(\cdot)$, si $d(\cdot)$ peut engendrer indéfiniment plus de profits que la “*buy and hold*”. En ce sens, ce chapitre est aussi une application en finance de la théorie de la calculabilité, qui pourrait intéresser des chercheurs

3. Les suites aléatoires au sens de Schnorr (1971) sont définies dans la section 3.3.2 page 108.

des deux disciplines ici rassemblées, finance et informatique.

Au lieu de comparer des stratégies selon leur espérance de gains, comme ce qui est souvent réalisé en finance, notre définition de cours imbattables s'inscrit dans un nouveau cadre théorique : celui de la théorie de la calculabilité.

Traditionnellement, une série financière est toujours présentée dans le cadre des probabilités. Par exemple, en supposant que la série de rentabilités x_0, x_1, \dots, x_n est réalisée par une collection de variables aléatoires *i.i.d.* $N(0, 1)$, on déduit les propriétés de x_0, x_1, \dots, x_n en étudiant cette dernière collection de variables. Comme celle-ci peut engendrer un nombre infini de suites différentes, en réalité, on analyse toujours **une famille** de séries probables mais non une suite spécifique. L'étude d'un échantillon réel dans le cadre des probabilités consiste donc à identifier "la population" à laquelle celui-ci peut bien appartenir. Les conclusions obtenues sont toujours de type " x_0, x_1, \dots, x_n est réalisée par une collection de variables aléatoires *i.i.d.* $N(0, 1)$ qui est imbattable".

En revanche, dans le cadre de la calculabilité, on peut étudier directement une série spécifique, et formuler des conclusions comme "**la** série x_0, x_1, \dots, x_n est imbattable". C'est justement cette nouveauté conceptuelle qui nous permet d'appliquer les outils algorithmiques sur des séries financières.

Principaux résultats obtenus

À l'issue de cette étude sur les régularités financières, 3 résultats nous semblent devoir être mis en avant :

- (a) Les trois théories financières étudiées admettent (ou prédisent) l'imprévisibilité des séries boursières. Plus précisément, l'absence de rentabilités excessives est une conséquence de l'efficacité informationnelle parfaitement compatible avec la finance comportementale ou les approches conventionnalistes de la finance.

En revanche, à défaut d'un modèle d'équilibre universel, l'identification des rentabilités excessives reste équivoque. C'est pourquoi l'impossibilité de "battre le marché" prend un rôle prédominant dans les tests de l'efficacité informationnelle.

- (b) Ce travail contribue également à mieux définir les termes "battre le marché". Contrairement à l'usage courant, "battre le marché" au sens de l'efficacité est équivalent à l'impossibilité de battre la "*buy and hold*" à long terme. Par ailleurs, la dominance du portefeuille de marché n'est qu'une conséquence du CAPM, lui même modèle d'équilibre souvent remis en question par les faits empiriques.

Vérifiée d'abord par des tests statistiques, puis par des algorithmes de compression, l'impossibilité de battre la stratégie "*buy and hold*" est confirmée

sur le plan empirique, ce qui est sans doute un argument en faveur de la théorie de l'efficience.

En proposant une modélisation mathématique des cours imbattables, on pourra étudier l'impact de différentes politiques d'échange sur l'ensemble des dynamiques boursières. Par exemple,

- En étudiant l'impact de la politique de blocage sur la performance des stratégies calculables, on propose un argument en faveur des *“stop loss rules”* : si cette pratique n'améliore pas la performance des stratégies considérées, elle n'est pas non plus responsable de leur échec à long terme.
- Après avoir introduit les stratégies généralisée - celles qui peuvent varier leur mise à chaque fois qu'elles prennent une décision - on démontre que celles-ci peuvent battre des séries de cours qui sont imbattables pour des stratégies *“tout acheter et tout vendre”*.
- La vitesse à laquelle une stratégie engendre des différences de gains n'a pas d'incidence sur sa performance finale. La vitesse et la performance en gains sont deux critères relativement indépendants pour évaluer une stratégie donnée.

- (c) Enfin ce travail introduit la notion de la complexité de Kolmogorov en finance. En appliquant l'approche algorithmique aux variations journalières de l'indice Dow Jones, on montre qu'un degré d'aléa très élevé domine ces données journalières, ce qui confirme dans une certaine mesure, l'impossibilité de battre la *“buy and hold”* sur l'indice américain étudié. Bien que les données *tick-by-tick* s'avèrent plus compressibles que les journalières, en comparant différents marchés internationaux, on conclut que l'origine d'une telle compression se trouve probablement dans des facteurs micro-structuraux. Toutefois, ces taux de compression ne vont pas nécessairement de pair avec la possibilité de construire des stratégies engendrant des rentabilités excessives.

Bien que les algorithmes de compression n'aient pas détecté de régularités *“révolutionnaires”* permettant de battre le marché, l'introduction de l'approche algorithmique élargit sans doute la gamme des structures détectables, et, ce faisant, rend plus robuste notre conclusion sur l'impossibilité de battre la *“buy and hold”*. De plus, en proposant un indicateur d'aléa pour chaque série étudiée, la méthode algorithmique ouvre de nouvelles perspectives dans les comparaisons des différents marchés financiers.

1. HASARD, STRUCTURE ET PRÉVISIBILITÉ : UN POINT SUR LES THÉORIES EN FINANCE

Résumé

Comment les cours boursiers évoluent-ils ? Tel est le sujet de ce premier essai autour des régularités et imprévisibilités financières. L'imprévisibilité des rentabilités excessives, souvent considérée comme une implication de la théorie de l'efficience, est en réalité compatible avec les trois théories étudiées, i.e., i) l'efficience informationnelle, ii) la finance comportementale et iii) l'approche conventionnaliste. Car si les "surprises" en informations fondamentales sont imprévisibles par définition, il en va de même pour les changements d'émotion, d'humeur ou d'opinion des investisseurs. La rationalité des agents, qui se traduit dans l'approche classique par l'usage exclusif des informations fondamentales, devient en effet le seul point distinct des trois approches analysées.

Si le problème de l'hypothèse jointe compromet constamment les résultats des tests d'efficience, les investisseurs n'ont aucune raison valable de tenir les titres aux rentabilités négatives. Cette non-négativité des rentabilités espérées implique l'impossibilité de surperformer la "*buy and hold*" avec des stratégies du type "*market timing*".

En revanche, l'inefficacité des stratégies de "*stock picking*" est fondée sur le CAPM, valable uniquement dans un monde représenté par l'espérance et l'écart-type. En insistant sur cette différence entre l'efficience et le CAPM, les principales anomalies de marché sont étudiées dans la dernière section de ce chapitre. Compatibles avec l'impossibilité de "*market timing*", ces faits empiriques reviennent à des anomalies du CAPM parfaitement compatibles avec la théorie de l'efficience.

Introduction

Dans la littérature en finance, nous constatons trois théories concurrentes qui modélisent le comportement des investisseurs et qui impliquent des dynamiques boursières spécifiques :

1. la théorie de l'efficience, souvent considérée comme la représentante de l'approche néo-classique en finance (Samuelson 1965, Fama 1970, Lucas 1978, Ross 2004).
2. la finance comportementale qui explique les phénomènes financiers avec des concepts psychologiques (Daniel, Hirshleifer, et Subrahmanyam 1998, Shiller 2000),
3. enfin, la théorie conventionnaliste qui adopte une rationalité "Keynésienne" dans la perception des investisseurs (Orléan 1999).

Parmi les nombreux débats théoriques, l'imprévisibilité des rentabilités excessives (Fama 1970), ou l'absence de rentabilité excessive sur le long terme, est souvent considérée comme une implication directe de la théorie de l'efficience. Ici, le terme "rentabilité excessive" désigne la différence entre la rentabilité réalisée et la rentabilité d'équilibre. Elle est dite imprévisible (ou absente) si son espérance conditionnelle est toujours égale à zéro.

Or, dans ce chapitre, un point de vue élargi sera défendu au sujet de cette implication empirique :

"Quel que soit le courant théorique adopté, les rentabilités excessives des actifs financiers apparaissent toujours comme imprévisibles ; cette imprévisibilité découle de la libre rencontre entre les offres et les demandes sur les marchés financiers, inconditionnée à la rationalité des agents."

Selon la théorie de l'efficience, l'innovation des prix dépend uniquement des changements en informations fondamentales. Par "informations fondamentales" on entend les informations économiques ou financières qui ont un effet sur les dividendes futurs d'un titre donné. La valeur d'un titre correctement estimée avec les informations fondamentales est ce qu'on appelle "une valeur fondamentale ou intrinsèque". La survenue des informations fondamentales étant par nature imprévisible, les rentabilités excessives sont imprévisibles sur un marché efficient.

Selon les défenseurs de la finance comportementale, les variations des prix sont aussi bien influencées par les informations fondamentales que par les facteurs psychologiques des agents. Certains d'entre eux n'hésitent pas à affirmer la possibilité de prévoir des rentabilités excessives en intégrant des apports de la psychologie en finance (DeBondt et Richard 1985, Shefrin 2005).

Or, si le recours aux facteurs psychologiques est le point commun des travaux en finance comportementale, l'opinion diverge concernant la prévisibilité des rentabilités excessives. En effet, beaucoup d'auteurs (DeLong, Shleifer, Summers, et Waldmann 1990, Balduzzi, Bertola, et Foresi 1995, Brav et Heaton 2002, Shiller 2003, Chiarella, Dieci, et He 2007) considèrent les erreurs psychologiques comme une composante aléatoire de la rentabilité, car celles-ci ne devraient pas être plus prévisibles que les informations fondamentales.

En effet, si les erreurs psychologiques peuvent être identifiées de façon **ex post** dans les observations historiques, la variété des modèles comportementaux rend souvent difficile la formulation des "prédictions réfutables qui captent mieux la carte des anomalies que l'efficience du marché (Fama 1998, p.291)".

En proposant une rationalité stratégique, la finance conventionnaliste soutient également l'insuffisance des informations fondamentales dans l'explication des dynamiques boursières. Or, selon cette approche, si les investisseurs laissent survenir un écart entre le prix d'un actif et sa valeur fondamentale, c'est bien pour maximiser leurs intérêts financiers, mais non à cause des erreurs psychologiques totalement inconscientes. Considérées comme la conséquence des interactions stratégiques menées par l'ensemble des investisseurs, les rentabilités excessives ne sont pas plus faciles à prédire que pour les deux approches précédentes, car cela revient à anticiper les changements de l'opinion moyenne sur le marché (les ruptures de convention) dont les causes sont souvent ignorées par les conventionnalistes.

Comme l'imprévisibilité des rentabilités excessives est une implication commune des trois courants théoriques cités ci-dessus, tenter de réfuter la théorie de l'efficience en calibrant les séries des rentabilités successives ne semble ni une direction de recherche aisée ni prometteuse.

En revanche, l'efficience informationnelle, comme toute hypothèse scientifique, n'est pas infalsifiable. La rationalité des investisseurs peut différencier l'efficience des deux autres courants théoriques. Il sera défendu, dans ce chapitre, que seules les études empiriques mettant en évidence la rationalité (ou irrationalité) des marchés financiers pourraient tester rigoureusement l'efficience des marchés contre les deux autres approches théoriques.

Or, pour vérifier si les innovations de prix contiennent des erreurs d'évaluation, il faut d'abord se doter d'un modèle d'équilibre calculant les rentabilités qu'on devrait observer sous l'hypothèse de la rationalité. C'est justement sur le choix de ce modèle que les opinions divergent parmi les chercheurs en finance. Par conséquent, tout test de rationalité est accompagné d'une hypothèse supplémentaire sur le modèle d'équilibre utilisé. Ce problème connu sous le nom *d'hypothèse jointe* se trouve à l'origine du paysage obscur autour de l'efficience

informationnelle.

Si de considérables difficultés conceptuelles sont encore à franchir pour distinguer l'efficacité informationnelle de ces théories concurrentes, la validité de la loi de l'offre et de la demande dispose d'une implication directement testable : **“face à une série de cours donnée, aucune stratégie active ne peut gagner contre la stratégie “buy and hold”, ou la stratégie qui achète dès le premier instant et revend au dernier.”** En rejetant cette proposition, on pourrait rejeter l'hypothèse d'efficacité sans pour autant valider les approches concurrentes. Ce chapitre insiste sur les racines d'une telle proposition, et souligne son indépendance de toute hypothèse sur la rationalité des agents.

Ce premier essai sur l'imprévisibilité des rentabilités excessives amène donc à présenter un débat théorique autour de la nature (ordonnée ou aléatoire) des dynamiques de prix. Le but est de montrer, sur le plan conceptuel, que le hasard est en réalité une conséquence directe de la loi de l'offre et de la demande, et ce, quelle que soit la rationalité postulée pour les agents.

Ce chapitre est organisé en 3 parties :

1. Les trois courants théoriques cités précédemment seront analysés respectivement dans l'objectif de montrer leur implication sur la nature aléatoire des rentabilités excessives.
2. Une étude sur les modèles d'équilibre présentera en détail le problème de l'hypothèse jointe ainsi que ses implications sur les tests d'efficacité informationnelle.
3. Les anomalies du marché seront enfin analysées pour montrer leur compatibilité avec la nature imprévisible des rentabilités excessives.

1.1 Imprévisibilité des rentabilités excessives

Un marché concurrentiel est un marché où les acteurs ne sont soumis à aucune entrave d'entrée ou de sortie, où les offreurs et les demandeurs ne peuvent influencer la cotation des biens, et où les informations concernant les volumes et les prix de transactions sont disponibles instantanément à tous les participants. Étant données les conditions d'échanges des marchés financiers modernes, ces derniers sont souvent considérés comme de bonnes approximations des marchés concurrentiels où on constate une libre rencontre entre les offres et les demandes des actifs cotés.

Ainsi, sur un marché financier, le prix d'un actif augmente (resp. baisse) quand

la demande est supérieure (resp. inférieure) à l'offre. Celui-ci se stabilisera à un niveau qui égalisera l'offre et la demande de cet actif sur le marché.

Il est montré dans cette section que ce mécanisme de marché domine les différentes spécifications de la rationalité des investisseurs. Il offre un cadre très général pour les analyses théoriques des dynamiques boursières. À l'aide de ce cadre général, on montera les principaux rapprochements et divergences des courants théoriques au regard de leurs implications sur la prévisibilité des rentabilités excessives.

Le point de vue ici défendu est le suivant : ce qui distingue les principales théories actuelles réside dans leurs analyses de la rationalité, et aucune ne propose de méthode efficace permettant de prévoir les rentabilités excessives.

1.1.1 Rentabilité espérée, anticipée, exigée et d'équilibre, les moteurs du mécanisme de marché.

“Rentabilité espérée”, “rentabilité anticipée”, “rentabilité exigée”, et “rentabilité d'équilibre” sont des expressions souvent rencontrées dans la littérature en finance. Bien que leur signification économique soit sensiblement distincte, ces expressions sont souvent rassemblées dans une même écriture mathématique, celle de l'espérance conditionnelle, notée $E(\cdot|\cdot)$. Pour éviter toute confusion possible, il est nécessaire de revenir sur les définitions respectives de ces quatre expressions économiques.

Une telle précision est fondée sur les différences entre le prix espéré, le prix anticipé et la valeur fondamentale.

- Le prix espéré, noté $E(p_{t+1}|\Omega_t)$, est l'espérance conditionnelle du prix p_{t+1} à l'instant t . Par définition, la différence entre cette espérance **ex ante** et la réalisation **ex post** de p_{t+1} est imprévisible dans la mesure où son espérance conditionnelle est toujours nulle.
- En déployant uniquement les informations fondamentales à l'instant t , notées Ω_t^* , il est aussi possible de calculer l'espérance de la valeur fondamentale $E(p_{t+1}^*|\Omega_t^*)$. Bien que le choix du modèle pour calculer une telle valeur reste encore à identifier, on suppose ici l'existence de la valeur fondamentale.
- Le prix anticipé, noté $p_{t+1}^m(\Omega_t)$, est la croyance des investisseurs sur le prix futur p_{t+1} . Ici, $p_{t+1}^m(\Omega_t)$ désigne la fonction adoptée par les investisseurs dans leur processus d'anticipation. Selon la rationalité qu'on impose aux agents économiques, ce prix anticipé peut s'avérer différent de la valeur fondamentale $E(p_{t+1}^*|\Omega_t^*)$ ¹.

1. Comme Beaver (1981) l'a montré, l'existence de la fonction $p_{t+1}^m(\Omega_t)$ n'implique pas l'ho-

La croyance des investisseurs sur le prix futur, $p_{t+1}^m(\Omega_t)$, peut différer de sa vraie espérance $E(p_{t+1}|\Omega_t)$, car les individus peuvent se tromper en utilisant des mauvais modèles, en se laissant influencer par des rumeurs, etc. L'écart entre $p_{t+1}^m(\Omega_t)$ et $E(p_{t+1}|\Omega_t)$ est à notre sens **une erreur d'anticipation**.

Par ailleurs, le prix anticipé pourrait, selon certains, différer de la valeur fondamentale, c'est là un point débattu dans la littérature. L'écart entre $E(p_{t+1}^*|\Omega_t^*)$ et $p_{t+1}^m(\Omega_t)$ est **une erreur d'évaluation**.

Après l'introduction des prix espéré et anticipé, on peut décrire le mécanisme de marché qui permet de fixer le prix à l'instant t , p_t :

- Pour un titre donné, à chaque instant t , les investisseurs fixent une rentabilité exigée, r_{t+1}^* , en fonction des informations disponibles sur le marché, notamment la rémunération des actifs sans risque et le risque couru. Cette **rentabilité exigée** représente la récompense revendiquée par les agents pour la détention du titre durant la période $t \rightarrow t + 1$. C'est cette rentabilité qui est centrale dans le CAPM.
- La **rentabilité espérée** est calculée par $E(r_{t+1}|\Omega_t) = \frac{E(p_{t+1}|\Omega_t) - p_t}{p_t}$.
- La **rentabilité anticipée** est définie par $r_{t+1}^m = \frac{p_{t+1}^m(\Omega_t) - p_t}{p_t}$.
- Si $r_{t+1}^m > r_{t+1}^*$, les investisseurs achètent pour profiter de la rentabilité anticipée. Ces achats feront augmenter le prix p_t , ce qui réduira r_{t+1}^m . Dans le cas contraire, les ventes feront baisser p_t , ce qui augmentera r_{t+1}^m .

Le marché atteint son équilibre à l'instant t quand p_t est fixé à un niveau tel que $r_{t+1}^m = r_{t+1}^*$. C'est pour cela que la **rentabilité exigée** r_{t+1}^* est aussi nommée la **rentabilité d'équilibre**.

Ce mécanisme est d'ailleurs retenu par Fama (1976) pour défendre sa définition de l'efficience informationnelle :

“The characteristics of the market assessed distribution ($p_{t+1}^m(\Omega_t)$ selon notre notation) determine the equilibrium expected returns on securities (r_{t+1}^ selon notre notation), and the market then sets the prices of securities at t so that it perceives expected returns to be equal to their equilibrium values.” (Fama 1976, p.143)*

Néanmoins, la libre rencontre entre les offres et les demandes à notre sens n'implique pas l'efficience informationnelle : l'égalité entre “ $p_{t+1}^m(\Omega_t)$ ”, “ $E(p_{t+1}|\Omega_t)$ ” et “ $E(p_{t+1}^*|\Omega_t^*)$ ” n'est pas supposée par le mécanisme du marché. C'est uniquement sur un marché efficient que l'anticipation du marché doit être fondée sur les informations fondamentales et égale à la valeur fondamentale (Fama 1965).

Sans cette condition stricte sur la rationalité, le mécanisme du marché devient un cadre d'analyse qui s'applique à tous les marchés concurrentiels. Ce cadre

moogénéité des agents économiques. Si chaque investisseur utilise son propre modèle d'anticipation, alors, $p_{t+1}^m(\Omega_t)$ représenterait la moyenne pondérée des anticipations individuelles.

général sera utilisé dans la prochaine section, pour étudier les implications des trois principales théories financières sur la dynamique des séries boursières.

1.1.2 De l'efficacité informationnelle à la finance conventionnaliste : peut-on prévoir les rentabilités excessives ?

Dans cette section, on détaille comment le mécanisme du marché peut être utilisé pour étudier les trois théories concurrentes en finance : efficacité informationnelle, finance comportementale et finance conventionnelle, notamment au regard de leurs implications sur les dynamiques des séries boursières.

L'objectif d'une telle analyse est de défendre l'imprévisibilité des rentabilités excessives quel que soit le modèle de la rationalité retenu.

I. Efficacité informationnelle : anticipation rationnelle sur les marchés financiers

L'efficacité informationnelle est un terme utilisé par Fama (1970) pour décrire un état de marché où *“les prix reflètent pleinement et à tout instant l'information disponible.”* (Fama 1970, p.383)

Pour montrer les implications empiriques de cette définition, Fama (1976) lie l'efficacité informationnelle avec l'anticipation rationnelle de Muth (1961).

Selon Fama, sur un marché efficient, la rentabilité anticipée, r_{t+1}^m , est toujours égale à la rentabilité espérée, $E(r_{t+1}|\Omega_t)$, qui est équivalente à la rentabilité d'équilibre.

Les rentabilités successives sont donc modélisées par :

$$r_{t+1} = r_{t+1}^* + \varepsilon_{t+1} \tag{1.1}$$

Où ε_{t+1} satisfait $E(\varepsilon_{t+1}|\Omega_t) = 0$, et est nommé jeu équitable.

La rentabilité réalisée d'un actif donné est égale à la rentabilité exigée plus une erreur de réalisation. La rentabilité excessive, définie comme $r_{t+1} - r_t^*$, est donc imprévisible dans la mesure où son espérance conditionnelle est toujours égale à 0.

Ici, comme Samuelson (1965) et Fama (1970) le soulignent, sur un marché efficient, c'est la **rentabilité excessive** qui est un jeu équitable, mais non la

rentabilité réalisée. Selon l'équation (1.1), pour que les **rentabilités réalisées** soient uncorrélées, il faut supposer arbitrairement, la stabilité de la rentabilité exigée, r_{t+1}^* .

Or, Lucas (1978) et LeRoy (1989) ont démontré respectivement qu'une telle stabilité n'est pas nécessaire pour qu'un marché soit efficient. En effet, il est même plus rationnel pour les investisseurs d'exiger une récompense qui évolue selon les conditions économiques, notamment le risque du marché ou la rentabilité sans risque.

Si l'obtention du jeu équitable sous efficacité informationnelle s'établit de manière directe, l'interprétation du prix espéré, $E(p_{t+1}|\Omega_t)$, nécessite davantage de précisions. Selon Fama (1965), sur un marché efficient, le prix d'un actif donné est toujours distribué autour de sa valeur fondamentale, et la valeur fondamentale d'un actif donné peut être interprétée de deux manières différentes :

“First, perhaps they just represent market conventions for evaluating the worth of a security by relating it with various factors which affect the earnings of a company. On the other hand, intrinsic value may actually represent equilibrium prices in the economic sense, i.e., prices that evolve from some dynamic general equilibrium model.”
(Fama 1965, p.36)

En définissant la rentabilité exigée r_{t+1}^* à l'aide d'un modèle d'équilibre général comme celui de Lucas (1978), l'efficacité informationnelle prend un sens économique très particulier : sur un marché efficient, l'allocation des ressources économiques est toujours optimale. L'équilibre de marché, sous cet angle, est un optimum de Pareto où l'amélioration du bien-être d'un individu n'est réalisable qu'en détériorant celui des autres.

Or, en décalage avec l'approche néo-classique qui stipule l'optimum de Pareto, Fama ne lie pas l'efficacité informationnelle à cette dernière propriété. Selon cet auteur, la valeur intrinsèque d'un actif peut aussi émerger “d'une convention partagée sur le marché” qui peut parfaitement s'écarter des modèles d'équilibre général.

La seule condition que Fama semble vouloir imposer à ces conventions de marché se trouve dans la rationalité des agents : les informations prises en compte par le marché doivent être de nature fondamentale. Cette condition est indispensable pour différencier l'efficacité informationnelle de la finance comportementale ou conventionnaliste, car si les conventions de marché peuvent aussi bien découler des facteurs psychologiques, alors l'effet de ces derniers sur les prix agrégés ne serait plus une preuve d'inefficacité. Le débat actuel entre les différentes approches financières ne disposerait plus de support théorique.

Sur un marché efficient, la rentabilité d'équilibre, r_{t+1}^* , et la vraie distribution du prix futur, $E(p_{t+1}|\Omega_t)$, sont uniquement influencées par les informations fondamentales. L'intégration des facteurs non fondamentaux dans ces estimations serait une preuve d'inefficience. Ce lien entre l'efficience informationnelle et les informations fondamentales est probablement la seule condition définitive sur la rationalité de Fama (1970).

Dans ses premières formulations, deux autres conditions (information parfaite et hétérogénéité des agents) sont souvent associées la théorie de l'efficience, et ce, malgré la précision de Fama (1970) citée ci-dessous :

“By frictionless market in which all information is freely available and investors agree on its implications is, of course, not descriptive of markets in practice. Fortunately, these conditions are sufficient for market efficiency, but not necessary. For example, as long as transactors take account of all available information, even large transaction costs that inhibit the flow of transactions do not in themselves imply that when transactions do take place, prices will not “fully reflect” available information. Similarly, the market may be efficient if “sufficient numbers” of investors have already access to available information”. And disagreement among investors about the implications of given information does not in itself imply market inefficiency unless there are investors who can consistently make better evaluations of available information that are implicit in market prices.” (Fama 1970, p.388)

Suite à cette légère confusion entre les marchés efficient et parfait, les améliorations de l'efficience informationnelle consistent souvent à la reformuler dans un cadre d'information imparfaite et/ou d'agents hétérogènes (Grossman et Stiglitz 1980, Beaver 1981, Latham 1986) (l'annexe A récapitule les principaux travaux dédiés à ce sujet). Cependant, si ces débats ultérieurs ont pu apporter un éclairage sur l'interprétation de la rationalité prêtée aux investisseurs, l'implication de l'efficience informationnelle sur les rentabilités successives (exprimée par l'équation 1.1) reste au coeur de la théorie financière 40 ans après sa formulation initiale (Ball 1989).

Vérifier si l'équation (1.1) rend fidèlement compte des séries de rentabilités observées est le principe de base pour tous les tests d'efficience.

Or, l'imprécision du modèle d'équilibre permettant de calculer r_{t+1}^* , risque de compromettre toute conclusion tirée des tests empiriques :

“Market efficiency says nothing specific about how the 'true distribution of p_{t+1} ' determines the equilibrium expected return. This is the

province of the model of market equilibrium. Some model of market equilibrium, that is, some specific statement about how 'expected return' is related to the 'true distribution of p_{t+1} ' is needed to test the market efficiency condition, but the choice of a model of equilibrium is not restricted by the market efficiency condition.” (Fama 1976, P.144)

Ce problème d'hypothèse jointe rend les éventuels écarts entre les séries de rentabilités observées et l'équation (1.1) difficiles à interpréter. Ceci risque de limiter considérablement la réfutabilité, et donc la nature scientifique, de l'hypothèse d'efficience. La discussion à ce sujet est reportée dans la section 1.1.3 (page 27).

Pour résumer, trois propositions principales concernant la théorie de l'efficience sont mises en avant dans cette section :

1. un marché efficient n'a pas besoin d'être parfait, il n'implique pas toujours l'homogénéité des agents économiques ou la gratuité des informations. Le seul point qui distingue la théorie de l'efficience des approches concurrentes est la rationalité des agents économiques, c'est-à-dire l'attachement des rentabilités d'équilibre aux facteurs qu'ils jugent fondamentaux ;
2. d'un point de vue empirique, sur un marché efficient, les rentabilités peuvent être modélisées par l'équation (1.1). La rentabilité d'équilibre émerge grâce aux interactions entre les investisseurs ;
3. faute d'accord sur le choix du modèle d'équilibre, l'efficience informationnelle est difficilement testable sur le plan empirique.

Si l'imprévisibilité des rentabilités excessives est une implication admise de la théorie de l'efficience, les opinions divergent quant à la relation entre cette imprévisibilité et la finance comportementale. Une discussion à ce sujet est organisée dans la section suivante pour défendre l'impossibilité de prévoir les erreurs d'anticipation même selon l'approche comportementale.

II. Finance comportementale : introduction des facteurs psychologiques

On peut relier la naissance du courant dit de la finance comportementale à l'article de Burell (1951) intitulé *“Possibility of an Experimental Approach to Investment Studies”*, dans lequel l'auteur, professeur à l'université d'Oregon, a fait un premier usage des expériences de laboratoires en économie pour vérifier des hypothèses financières.

Ce travail séminal a été suivi, dans les années 70 par les chercheurs en psychologie (Slovic 1972, Slovic et Bauman 1972, Tversky et Kahneman 1974), pendant que les économistes financiers se sont principalement concentrés sur le débat autour d'efficience. C'est uniquement à partir des années 80 (Shiller 1981, DeBondt

et Richard 1985), après les premières constatations des anomalies du marché, que la finance comportementale a commencé à prendre du terrain vis-à-vis de la théorie de l'efficience.

D'une manière générale, la finance comportementale est définie par l'introduction des biais psychologiques dans les interprétations des phénomènes financiers.

Parmi d'autres, on peut évoquer :

1. "*Over-confidence*" (excès de confiance) : les investisseurs sont souvent trop confiants en leurs jugements personnels, cela les entraîne à surestimer le poids des informations privées dans leurs modèles d'anticipation (Bem 1965, Alpert et Raiffa 1982, Daniel, Hirshleifer, et Subrahmanyam 1998).
2. "*Conservatism*" (conservatisme) : les agents ont une tendance à prédire le futur en se basant sur les observations passées. Ce biais les empêche d'actualiser leur modèle de prédiction face aux informations récentes (Edwards 1968, Andreassen et Kraus 1990).
3. "*Representativeness heuristic*" (routines mentales ou heuristiques) : les individus sont souvent trop concentrés sur les "tendances" artificielles pour se rappeler de la vraie population des données analysées (Kahneman et Tversky 1979, DeBondt et Richard 1985).
4. "*Safety first principal*" (aversion à la perte) : les investisseurs ont beaucoup plus de mal à accepter une perte qu'un bénéfice. Ce biais les pousse à surréagir aux informations négatives (Shefrin et Statman 2000).

Ces biais psychologiques sont souvent combinés pour expliquer l'évolution des cours boursiers. Trois modèles comportementaux (i.e Barberis, Shleifer, et Vishny (1998), Daniel, Hirshleifer, et Subrahmanyam (1998), Hong et Stein (1999)) sont présentés dans l'annexe D pour illustrer l'usage de ces biais psychologiques dans l'explication des rentabilités excessives.

Étant donnés ces biais comportementaux, les investisseurs ne sont pas en mesure d'anticiper la valeur fondamentale des titres. L'équivalence entre le prix anticipé, le prix espéré et la valeur fondamentale que nous pouvons constater sur un marché efficient est donc remise en question par l'approche comportementale. Les erreurs d'évaluation provoquées pour des raisons psychologiques semblent donc ouvrir une porte à la prédiction des rentabilités excessives.

Or, si l'introduction des biais comportementaux sont de nature à améliorer nos compréhensions des faits empiriques, l'abondance des biais identifiés et la relation complexe entre ces biais et des décisions économiques réduisent la prévisibilité des erreurs d'évaluation. Comme Rick et Loewenstein (2008) l'ont précisé dans leur livre "*Handbook of Emotions*", des raisonnements économiques du type "*bad weather should lead to negative emotions which should in turn lead to negative price*

mouvements” sont trop simplifiés d’un point de vue psychologique, car “*negative emotions can exert conflicting effects on risk-taking* (Rick et Loewenstein 2008, p.153)”. L’impact des biais comportementaux sur les prix des actifs semblent être contextes-dépendants, ce qui rend difficile la formation de prédictions efficaces.

La finance présente de grandes différences avec les autres branches des sciences sociales, notamment la psychologie. Son objectif final n’est pas de décrire les investisseurs, mais de modéliser les conséquences de leurs comportements sur les prix agrégés. Chaque élément introduit dans les théories financières doit avoir un impact sur les prix du marché. Ce principe, inspiré du rasoir d’Occam, a été souligné par Lucas de manière explicite :

“A ‘theory’ is not a collection of assertions about the behavior of the actual economy but rather an explicit set of instructions for building a parallel or analogue system - a mechanical, imitation economy. A ‘good’ model, from this point of view, will not be exactly more ‘real’ than a poor one, but will provide better imitations. (Lucas 1980, p.697).”

Les biais psychologiques mis en avant par la finance comportementale couvrent une variété tellement large, et leur combinaison dans un modèle donné est tellement flexible, que la robustesse de leur effet sur les prix agrégés est souvent remise en question. Miller (1986) souligne ce point avec beaucoup d’humour :

“Behind each holding may be a story of family business, family quarrel, legacies received, divorce settlements and a host of other considerations... That we abstract from all these stories in building our models is not because the stories are uninteresting but because they may be too interesting and thereby distract us from the pervasive market forces that should be our principal concern (Miller 1986, p.467). “

Fama n’hésite pas à écrire :

“Given the demonstrated ingenuity of the theory branch of finance, and given the long litany of apparent judgment biases unearthed by cognitive psychologists (DeBondt and Thaler, 1995), it is safe to predict that we will soon see a menu of behavioral models that can be mixed and matched to explain specific anomalies. My view is that any new model should be judged on how it explains the big picture. The question should be : Does the new model produce rejectable predictions that capture the menu of anomalies better than market efficiency? For existing behavioral models, my answer to this question (perhaps predictably) is an emphatic no (Fama 1998, p.291).”

Rubinstein affirme que “la finance comportementale a en effet ouvert “la boîte de Pandore” remplie des histoires ad hoc qui ont peu de pouvoir de prédiction (Rubinstein 2001, p.16).”

Pour résumer, 3 arguments soutiennent l'imprévisibilité des rentabilités excessives en adoptant une approche comportementale :

- **Les modèles comportementaux ne quantifient pas l'envergure des erreurs d'évaluation.** De ce fait, selon les périodes de marché, la même information positive peut aussi bien provoquer une augmentation de 1% qu'une bulle de 100%.

Comme Leroy le note :

“we argued that attributing the stock price run-up to fundamentals is implausible in the absence of showing why apparently similar earlier trends in fundamentals did not produce similar dramatic effects on security prices. The argument, of course, applies equally to (irrational) bubble explanations : why a bubble develop in the 1990 not earlier ? (LeRoy 2004, p.802)”

En présence de coûts de transactions, la réalisation des gains anormaux (rentabilités excessives) nécessite de distinguer les erreurs infimes de celles qui sont suffisamment grandes à exploiter, ce qui ne facilite pas l'exploitation financière des biais psychologiques.

- **La frontière entre les tendances provoquées par les biais psychologiques et les retournements de situation sont extrêmement obscures.** À notre connaissance, aucun modèle comportemental n'est en mesure de prévoir le moment critique où les agents irrationnels abandonnent la tendance poursuivie et déclenchent un retour à la valeur fondamentale. Comme Hong et Stein (1999) l'ont précisé :

“In our model, momentum traders ... impose a negative externality on others, because someone entering the market at any time t does not know how heavily invested momentum traders are in the aggregate at this time, and hence cannot predict whether or not there will be large-scale unwinding of momentum positions in the near future (Hong et Stein 1999, 2167).

- **Distinguer les gains normaux de ceux qui ne le sont pas est une tâche extrêmement difficile, voire théoriquement impossible, et ce, pour 2 raisons distinctes :**

1. La mesure des rentabilités excessives, $r_{t+1} - r_{t+1}^*$, nécessite un modèle spécifique de la rentabilité d'équilibre r_{t+1}^* . Le problème de l'hypothèse jointe laisse toujours un doute sur la présence de cet écart (Fama 1990, Fama et French 1993, Fama et French 1996, Constantinides 2002, Boehme et Sorescu 2002).
2. Les phénomènes expliqués par les modèles comportementaux peuvent souvent l'être par les facteurs fondamentaux. Par exemple, selon Johnson (2002), la croissance des entreprises peut expliquer la plupart des phénomènes constatés dans les travaux empiriques et qui sont considérées comme les preuves des biais psychologiques par les comportementalistes.

Pour conclure, la finance comportementale en tant que théorie complémentaire de l'approche classique, propose des explications intéressantes aux phénomènes observés sur les marchés financiers. Or, quand il s'agit de la prévision et de l'exploitation des rentabilités excessives, cette théorie ne semble pas être suffisamment efficace.

Cette position semble coïncider avec celle de Shiller, pourtant partisan de la finance comportementale. Dans son article de 2003, l'auteur conclut :

“Of course, we do not expect such research to provide a method to make a lot of money out of financial market inefficiency very fast and reliable. We should not expect market efficiency to be so egregiously wrong that immediate profits should be continuously available. But market efficiency can be egregiously wrong in other senses. For example, efficient market theory may lead to drastically incorrect interpretations of events such as major stock market bubbles (Shiller 2003, p.101).”

III. Finance conventionnaliste : vers une rationalité stratégique

La finance conventionnaliste adopte une rationalité “Keynesienne” dans ses analyses des marchés financiers (Orléan 1999, Poitras 2002). Au contraire des deux approches précédentes, cette théorie se développe dans un cadre “d’incertitude absolue (Knight 1921, p.20)” où aucune prédiction sous forme de probabilité ne peut être formulée. L’anticipation rationnelle, au sens de Savage (1954) et de Muth (1961), est tout simplement hors de propos.

En raison de cette incertitude absolue, les prix de transactions ne peuvent qu’être de simples agrégations des opinions individuelles. Les investisseurs conscients et intelligents adoptent par conséquent une rationalité stratégique en soulignant le rôle de la coopération collective.

Comme la finance comportementale, l’approche conventionnaliste soutient aussi l’écart entre le prix d’un actif et sa valeur fondamentale. Cependant, le retour de valeur à très long terme², souvent mobilisé par la finance comportementale pour expliquer l’auto-corrélation négative des rentabilités décennales, est entièrement renié par les conventionnalistes. Le décrochage complet entre le prix des actifs et leur valeur fondamentale est en effet la principale divergence entre la finance comportementale et l’approche conventionnaliste.

2. Dans les travaux en finance comportementale, la valeur fondamentale est souvent considérée comme la limite du prix à l’horizon infini.

Dans cette section, nous allons étudier la rationalité stratégique se trouvant au coeur de la finance conventionnelle en essayant de mettre en évidence le processus de prix impliqué par cette rationalité, ainsi que les difficultés quasi-insurmontables pour prédire les rentabilités excessives.

L'anticipation rationnelle n'a pas sa place dans un monde incertain. Selon la théorie des probabilités, il existe une "vraie probabilité" pour chaque événement. Par exemple, quand on lance un dé équilibré, la "vraie probabilité" d'obtenir chacun de ses côtés est de $\frac{1}{6}$.

Or, la "vraie probabilité" d'un événement donné n'est pas toujours connue d'avance. Pour l'évaluer, une méthode d'estimation s'impose : en observant les résultats des expériences historiques on peut inférer les "vraies probabilités" des différents scénarios avec leur fréquence respective.

Identifier les scénarios possibles et observer leur fréquence d'apparition, tel est le principe des conjectures statistiques.

Or, le monde réel est beaucoup plus complexe qu'un dé équilibré. Estimer la fréquence d'un événement donné nécessite une classification efficace permettant de regrouper les expériences similaires³.

Toute erreur de classification rend absurdes des conjectures statistiques. Knight (1921, chap.VII) en donne un exemple : le taux de mortalité des soldats américains aux Philippines est moins élevé que celui de la population restant aux États-Unis, on peut donc conclure qu'il est moins dangereux de faire la guerre aux Philippines que vivre paisiblement aux États-Unis. Cette absurdité est due à une mauvaise classification des observations, car la mortalité n'est pas uniquement influencée par l'endroit où l'on vit et le métier que l'on mène, elle dépend avant tout de l'âge et de l'état de santé des individus observés. Les jeunes hommes en bonne santé envoyés aux Philippines présentent naturellement une mortalité moins élevée que la population totale des États-Unis.

La maîtrise d'une situation donnée dépend donc de la possibilité d'effectuer une classification fiable des scénarios éventuels. Knight distingue trois situations différentes⁴ :

1. Les "vraies probabilités" sont connues : l'archétype est ici le lancé d'un dé équilibré.

3. Le chapitre VII de Knight (1921) offre une description précise de la méthode de classification.

4. On retrouve, d'ailleurs, ce même raisonnement dans la thèse de Keynes publiée en 1921. Dans son travail doctoral, Keynes utilise le terme "poids d'un argument donné" ("*weight of argument*" (Keynes 1921, p.4)) pour désigner la fiabilité d'une méthode de classification.

2. Les “vraies probabilités” sont inconnues, mais il est possible de classer les expériences observées et en déduire une probabilité suffisamment fiable. On parle dans ce cas de figure “d’une situation risquée”.
3. Parfois, nous ne connaissons ni les “vraies probabilités” des scénarios possibles, ni leur méthode de classification. Nous sommes dans “une incertitude absolue”.

Selon Keynes, le monde économique se trouve dans une incertitude absolue, les expériences économiques étant inrenouvelables, aucune des répliques ne partagerait les mêmes contextes sociaux. Par exemple, une entreprise peut voir son bénéfice diminuer suite à un changement stratégique, financier, ou de la conjoncture sociale. La complexité de la société humaine et la forte variation des contextes économiques impliquent l’échec de toute classification statistique. En d’autres termes, les groupes homogènes d’événements économiques ne peuvent pas exister.

Ainsi, “nos connaissances sur les facteurs qui vont gouverner le rendement d’un actif dans plusieurs années sont habituellement très minces et souvent négligeables Keynes (1936, chap.12)⁵”, et “certaines conséquences futures ne pourraient être associées à aucune probabilité Keynes (1936, p.148-150)⁶.” La rationalité des investisseurs, selon Keynes, ne peut pas s’exprimer en termes probabilistes⁷.

C’est donc sur cette incertitude absolue qu’est fondée la finance conventionnaliste, qui propose une nouvelle perspective pour se saisir de la rationalité des investisseurs.

Face à une incertitude absolue, les investisseurs ne peuvent pas former d’anticipation rationnelle concernant les valeurs fondamentales des titres. Il est même plus raisonnable pour eux d’abandonner ces tentatives d’évaluation, car :

“It would be foolish, in forming our expectations, to attach great weight to matters which are very uncertain. It is reasonable, therefore, to be guided to a considerable degree by the facts about which we feel somewhat confident, even though they may be less decisively relevant to the issue than other facts about which our knowledge is vague and

5. Le texte original en anglais : “our knowledge of the factors which will govern the yield of an investment some years hence is usually very slight and often negligible.”

6. Le texte original en anglais : “some future consequences could have no probability ratios assigned to them.”

7. Il convient ici de remarquer que l’incertitude dans le sens de Keynes ne coïncide pas entièrement avec la rationalité limitée de Simon (1983), tant citée par les comportementalistes (Kahneman 2003). La rationalité limitée souligne la légitimité des impulsions émotionnelles, des heuristiques psychologiques, alors que l’incertitude absolue de Keynes, comme nous l’avons précisé, s’appuie uniquement sur la complexité du mode réel, dans cette incertitude, même les investisseurs sans biais psychologiques ne pourront pas prévoir le futur avec les probabilités fiables (Dunn 2001).

scanty (Keynes 1936, chap.12)."

Ainsi, les facteurs fondamentaux ne serviront plus comme seul et unique support des stratégies rationnelles, car "*l'information pertinente n'est plus définie objectivement*", et "*la cotation observée sur un marché financier n'est rien d'autre que l'expression de l'opinion moyenne des intervenants* (Orléan 1989, p.257)."

Selon ce principe, **anticiper les prix futurs revient à prévoir l'opinion moyenne de l'ensemble des intervenants du marché**. La rationalité dans l'incertitude prend la forme d'un jeu coopératif entre les investisseurs. Comme beaucoup de théoriciens l'ont souligné, la coordination dans ce cas de figure se procède souvent à l'aide d'une convention collective (Schelling 1960, Lewis 1969).

L'exemple de la "coupure de téléphone", donné par Batifoulier, Merchiers, et Urrutiaguer (2002), illustre explicitement le rôle de la convention dans un jeu coopératif.

La communication téléphonique entre Franck et Guillemette est interrompue, qui doit rappeler en premier ? Ce jeu de coordination offre 4 issues possibles :

1. Franck et Guillemette rappellent tous les deux. La ligne sonnerait occupée, gain nul pour chacun (case "A" dans le tableau 1.1).
2. Guillemette ne rappelle pas et Franck rappelle. La communication se rétablit, chaque agent obtient un gain de 1 (case "B" dans le tableau 1.1).
3. Guillemette rappelle et Franck ne rappelle pas. La communication se rétablit, chaque agent obtient un gain de 1 (case "C" dans le tableau 1.1).
4. Personne ne rappelle. Gain nul pour les deux (case "D" dans le tableau 1.1).

Sans accord préalable, il est difficile pour les deux joueurs de distinguer les solutions B et C. Cependant, en se référant respectivement à la règle "celui qui a appelé, rappelle", la coordination entre eux sera optimisée. Cette règle du jeu est en effet une "saillance" au sens de Schelling (1960), et une "convention" selon Lewis (1969).

Tab. 1.1: Un exemple du jeu de coordination : coupure de communication téléphonique

		Guillemette	
		Rappelle	Ne rappelle pas
Franck	Rappelle	(0,0) A	(1,1) B
	Ne rappelle pas	(1,1) C	(0,0) D

Sur un marché financier, une convention collective simplifie considérablement le travail de coordination dans la mesure où tous les investisseurs peuvent réagir de la même manière en sachant que leur solution sera choisie par la majorité. Ainsi, un investisseur “n’a pas besoin de perdre le sommeil en se demandant combien son investissement vaudra dans 10 ans (Keynes 1936, chap.12) ⁸.”

Comme le prix d’un actif est l’opinion moyenne des investisseurs, l’égalité entre $E(r_{t+1}|\Omega_t)$ et r_{t+1}^m est systématique sous la stabilité de la convention collective. La rentabilité excessive définie par $r_{t+1} - r_{t+1}^*$ sera sans aucun doute d’espérance nulle.

Bien sûr, l’égalité $E(r_{t+1}|\Omega_t) = r_{t+1}^m$ serait brisée en cas d’une rupture de convention pendant la période $t \rightarrow t + 1$. Cependant, les ruptures de convention sont tellement compliquées à expliquer que, à nos connaissances, aucun modèle conventionnaliste ne peut prévoir l’écart $E(r_{t+1}|\Omega_t) - r_{t+1}^m$.

Comme Keynes l’a précisé au chapitre 12 de “Théorie générale de l’emploi, de l’intérêt et de la monnaie”,

“A conventional valuation which is established as the outcome of the mass psychology of a large number of ignorant individuals is liable to change violently as the result of a sudden fluctuation of opinion due to factors which do not really make much difference to the prospective yield; since there will be no strong roots of conviction to hold it steady. In abnormal times in particular, when the hypothesis of an indefinite continuance of the existing state of affairs is less plausible than usual even though there are no express grounds to anticipate a definite change, the market will be subject to waves of optimistic and pessimistic sentiment, which are unreasoning and yet in a sense legitimate where no solid basis exists for a reasonable calculation.”
(Keynes 1936, chap12)

Par conséquent, l’imprévisibilité des ruptures de convention implique celle des rentabilités excessives. Comme pour les deux théories précédentes, selon l’approche

8. Texte original : “need not lose his sleep merely because he has not any notion what his investment will be worth ten years hence .”

conventionnaliste, les rentabilités successives contiennent aussi deux composantes indépendantes : la rentabilité exigée et un terme d'erreur toujours imprévisible.

1.1.3 L'efficience informationnelle, comment la tester ?

Les développements dans la section précédente ont pour objectif de mettre en évidence une seule et unique conclusion : les théories financières sont avant tout les **explications** des phénomènes financiers, en ce qui concerne la **prévision** des rentabilités excessives, aucune ne semble fournir d'outil efficace valable à long terme.

Si l'imprévisibilité des rentabilités excessives est un point commun des différentes théories financières, peut-on encore les distinguer sur le plan empirique ? L'efficience informationnelle est-elle scientifiquement testable ? Une discussion autour de ces questions est nécessaire à ce stade de l'analyse.

Dans la littérature en finance, on constate deux catégories de tests d'efficience : (1) les tests sur la prévisibilité des rentabilités excessives, (2) les tests sur la nature des informations intégrées dans les prix.

En étudiant le fondement théorique de ces deux catégories de tests, il sera montré ci-dessous que l'efficience informationnelle est théoriquement réfutable, mais à cause du problème d'hypothèse jointe, les résultats obtenus sont difficiles à interpréter.

- L'imprévisibilité des rentabilités excessives, est une implication du mécanisme de marché. Cette contrainte, qui par ailleurs gouverne le "principe de non-arbitrage" mis en avant par certains auteurs (Malkiel 2003, Ross 2004), peut être valable sans formuler une hypothèse de rationalité contraignante. Par conséquent, les tests dans cette direction ne peuvent pas départager les trois théories concurrentes. Autrement dit, le rejet d'une telle imprévisibilité réfuterait la théorie de l'efficience sans pouvoir valider le rôle de la psychologie ou de la convention dans les prix agrégés. Un autre défaut des tests de prévisibilité se trouve dans la nature **ex post** des données utilisées. **Expliquer** la rentabilité observée r_{t+1} avec les informations observées à la période $t + 1$ est une chose, **prévoir** r_{t+1} avec les informations disponibles à la période t en est une autre. Un modèle qui explique les rentabilités excessives en mobilisant les éléments imprévisibles ne peut pas rejeter l'efficience des marchés financiers.
- Les facteurs explicatifs des rentabilités successives résident au coeur de la rationalité. En établissant des preuves tangibles sur le rôle de la psychologie dans les innovations des prix, on obtiendrait une réfutation de l'efficience

1.2. Absence de consensus sur le modèle d'équilibre, une source de confusion

et une confirmation des approches comportementale ou conventionnaliste. Or, ici, il ne s'agit pas de montrer en quoi les facteurs psychologiques peuvent expliquer les variations de prix. On cherche à prouver scientifiquement que la partie des rentabilités expliquée par des facteurs psychologiques ne peut pas l'être en mobilisant des facteurs fondamentaux. Étant donnée la flexibilité dont on bénéficie dans le choix et l'interprétation des facteurs fondamentaux, ce débat en cours entre Fama, French, Thaler et DeBondt semble extrêmement difficile à trancher (Fama 1990, Fama et French 2000, Thaler 1999, DeBondt 2002, Brav et Heaton 2002).

Pour résumer, le débat entre les théories financières se pose avant tout en termes conceptuels. Leurs implications empiriques sont extrêmement difficiles à tester. Ce point de vue sera développé d'avantage dans la section 1.3 page 43, quand les anomalies empiriques de l'efficience seront abordées.

1.2 Absence de consensus sur le modèle d'équilibre, une source de confusion

Les modèles d'équilibre font partie de "*l'asset pricing theory*", qui, avec la théorie de l'efficience, constituent les deux piliers de la finance néo-classique (Ross 2002, p.129). Certains modèles d'équilibre, et surtout le CAPM (Sharpe 1964, Lintner 1965), sont considérés comme une implication directe de l'efficience, une certaine confusion concernant ces deux théories néo-classiques apparaît par ailleurs dans les travaux empiriques.

L'objectif de cette section est de brièvement présenter les principaux modèles d'équilibre dans la littérature, en questionnant leur validité empirique, ainsi que leur lien avec l'efficience informationnelle. La deuxième partie de la section vise à différencier l'efficience de marché et le CAPM. Il sera argumenté que cette distinction théorique a des implications importantes sur la mesure des performances des fonds d'investissement.

1.2.1 Du CAPM au modèle multi-facteurs, comment calculer la rentabilité exigée ?

Le "*Capital Asset Pricing Model*" (CAPM) introduit par Sharpe (1964) et Lintner (1965) est le premier modèle d'équilibre empiriquement testable. Son ancrage théorique profond ainsi que sa simplicité à mettre en pratique, ont fait du CAPM un des modèles les plus souvent utilisés par les professionnels.

La conception du CAPM commence par la distinction entre le risque systématique (ou risque de marché) et le risque spécifique. Lee et Lee (2006) définissent ces deux types de risques comme suit :

“Market risk is a risk attributable to common macroeconomic factors. Firm-specific risk reflects risk peculiar to an individual firm that is independent of market risk. (Lee et Lee 2006, p.189)”

Par exemple, en janvier 2008, la Société Générale a perdu 5 milliards euros à cause d'une opération litigieuse. Cet élément qui lui est spécifique ne doit pas influencer, par exemple, la cotation du group Total. Le risque associé à ce type d'événements est nommé risque spécifique. En revanche, la récession économique que nous vivons depuis le troisième trimestre de 2008 a un effet sur l'ensemble des marchés financiers et correspond bien à la notion de risque systémique.

Markowitz (1959) montre qu'en détenant plusieurs actifs en même temps, on peut diversifier le risque spécifique. Cette approche nommée la “théorie du portefeuille” est fondée sur deux postulats de base :

- La rentabilité future de chaque actif suit un processus stochastique qui peut être caractérisé par son espérance (rentabilité) et sa variance (risque).
- Pour un niveau de rentabilité souhaité, les investisseurs cherchent à minimiser leur risque.

Selon le critère “rentabilité-risque”, tous les portefeuilles possibles sont enveloppés dans l'ensemble délimité par la courbe *abc* dans la figure 1.1 page 30. On constate que pour un niveau de risque donné, les portefeuilles sur la courbe *aTb* (sans prolongement) sont les seuls à maximiser la rentabilité. Ils sont “efficients” dans le sens de Markowitz, car ils vérifient un ratio de rentabilité-risque optimal.

Dans les années 60, les techniques calculatoires étant plus limitées, la théorie du portefeuille nécessitait trop de calculs d'optimisation pour une mise en pratique générale.

Pour simplifier l'optimisation du portefeuille, Tobin (1958) introduit avec l'actif sans risque, un théorème de séparation dit "des deux fonds". Si tous les investisseurs cherchent à investir dans le meilleur portefeuille sur le marché, chacun d'eux doit détenir exactement le même portefeuille risqué. Comme l'ensemble des actifs est détenu dans n portefeuilles identiques, où n désigne le nombre total d'investisseurs, alors, le meilleur portefeuille ne peut qu'être celui du marché, i.e., un portefeuille vérifiant les principes suivants :

1. contenir tous les titres cotés sur un marché financier.
2. la proportion de chaque actif coté est fixée selon sa capitalisation boursière.

Par définition, le portefeuille ne subit aucun risque spécifique, il offre la meilleure parité “rentabilité/risque”.

Ainsi, en supposant l'existence d'un actif sans risque sur le marché (noté R_f), les portefeuilles efficients, i.e., ceux qui offrent la meilleure rentabilité à un niveau de risque donné, se trouvent sur la droite R_fT dans la figure 1.1. Cette droite nommée "la frontière efficiente" ("*capital market line*"), désigne un portefeuille optimal à chaque niveau de risque donné.

Les portefeuilles sur "la frontière efficiente" sont les combinaisons entre le portefeuille de marché et l'actif sans risque. La ligne R_fT est tangente à la courbe abc au point T , car dans le cas contraire, il existerait un investissement dans la surface abc qui offrirait un meilleur rapport rentabilité/risque que celui du marché. Pour cette même raison, la courbe Tei dans la figure 1.1 qui représente toutes les

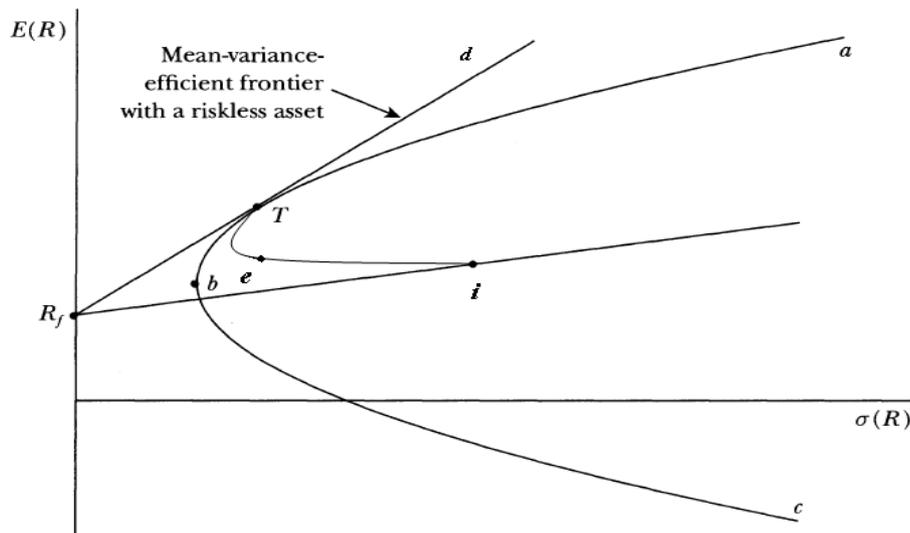


Fig. 1.1: Toutes les opportunités d'investissement.

L'axe X représente le niveau du risque mesuré en termes de l'écart-type, l'axe Y la rentabilité espérée. La surface à l'intérieur de la courbe abc représente tous les portefeuilles composables à partir des actifs risqués. L'exemple i est un portefeuille inefficent dans la mesure où celui-ci offre une relation rentabilité-risque moins intéressante que le portefeuille de marché. La ligne droite R_fT désigne les portefeuilles portant un ratio rentabilité/risque aussi élevé que celui du marché.

Source : reproduite à partir de la figure I de l'article Fama et French (2004).

combinaisons possibles entre un actif quelconque i et le portefeuille de marché (noté T), doit également être tangente à la ligne R_fT au point T . Au cas contraire, il existerait une combinaison entre i et T qui dépasse la frontière efficiente. Le portefeuille de marché n'offrirait plus le meilleur rapport rentabilité-risque.

Fondée sur ces relations de contingence, Sharpe (1964) et Lintner (1965) ont déduit la relation entre la rentabilité espérée d'un actif i et celle du marché m :

$$E_i = R_f + \beta_{i,m}(E_m - R_f) \quad (1.2)$$

Où E_i désigne la rentabilité espérée de i et E_m celle du marché. $\beta_{i,m} = \frac{\text{cov}(r_i, r_m)}{\sigma_m^2}$ mesure le risque systémique de l'actif i . La valeur de $\beta_{i,m}$ indique la variation moyenne de l'actif i , à chaque fois que le marché varie d'une unité. Par exemple, $\beta_{i,m} = 2$ signifie que l'actif i augmente (resp. baisse) de 2% en moyenne, à chaque fois que le marché augmente (resp. baisse) de 1%.

L'équation 1.2 est la formule de CAPM qui calcule pour chaque actif sa rentabilité d'équilibre. Pour vérifier la validité de cette formule, trois catégories de tests sont souvent envisagées :

1. Avec les données historiques, on peut estimer les paramètres du CAPM (R_f , $\beta_{i,m}$, ou $E_m - R_f$), puis comparer les valeurs estimées avec leurs valeurs théoriques (Blume 1970, Blume et Friend 1973, Fama et MacBeth 1973).
2. Il est possible aussi de composer des portefeuilles selon de différents principes, puis comparer leurs performances avec celui du marché. Le CAPM est rejeté si un des portefeuilles composés délivre une meilleure parité rentabilité/risque (Black, Jensen, et Scholes 1972, Jensen 1968).
3. On peut également tester l'effet des facteurs économiques sur les rentabilités financières. Le CAPM est rejeté s'il y a d'autres facteurs explicatifs que les termes de l'équation 1.2 (Fama et French 1993).

Ces trois catégories de tests seront présentées dans les sections suivantes où la validité du CAPM servira de fil de conducteur aux développements. D'autres modèles d'équilibre, proposés suite aux tests du CAPM, sont également analysés dans cette présentation. Il sera montré que malgré les efforts fournis dans cette direction, le modèle d'équilibre sans défaut reste encore à trouver.

I. Le "cross section test" et le CAPM de Black

Le "*cross section test*" introduit par Blume (1970) et développé par Fama et MacBeth (1973) est l'une des méthodes classiques pour vérifier les paramètres du CAPM. Le principe de ce test est assez intuitif : à l'aide d'une régression linéaire entre les β et E_i observés, on estime la prime de risque " $E_m - R_f$ ", et vérifie l'équivalence entre sa valeur estimée et théorique (c.f annexe B).

La plupart des "*cross section tests*" ont souligné une infériorité de la prime de risque estimée par rapport à sa valeur théorique. Pour expliquer cet écart, Black (1972) a proposé une version flexible du CAPM en mettant en avant l'accès limité aux emprunts obligataires.

En effet, dans la démonstration du CAPM, il est supposé que tous les investisseurs sont en mesure d'emprunter ou de prêter autant de capital qu'ils souhaiteraient au taux sans risque R_f . Étant données les conditions actuelles des marchés

financiers, cette hypothèse paraît peu réaliste. Selon Fisher Black, les investisseurs qui ne peuvent pas emprunter ou prêter au taux sans risque, remplaceraient l'actif sans risque par un actif risqué (noté z) indépendant du portefeuille de marché. z peut être des emprunts obligataires émis par une entreprise, un prêt de consommation individuel, etc.

En remplaçant R_f par E_z dans l'équation 1.2, la relation (1.3) doit être vérifiée entre un actif donné et le portefeuille de marché.

$$E_i = E_z + \beta_{i,m}(E_m - E_z) \quad (1.3)$$

Ici, E_z est la rentabilité espérée de z . Selon Black (1972) la seule condition sur E_z est que sa rentabilité espérée soit inférieure à celle du marché.

Cette version du CAPM, nommée "*Black-CAPM*", peut ainsi expliquer les résultats empiriques obtenus dans les "*cross section tests*". Remarquons que le CAPM classique et le "*Black CAPM*" s'accordent parfaitement sur la dominance du portefeuille de marché par rapport aux autres investissements possibles. Le CAPM sera donc réfuté s'il existe un portefeuille offrant un ratio rentabilité/risque plus élevé que celui du marché. L'existence d'un tel portefeuille est souvent vérifiée par les "*time series tests*".

II. Les "time series tests" et les performances des portefeuilles

Les tests en séries temporelles ("*time series tests*") introduits par Jensen (1968), Black, Jensen, et Scholes (1972) et développés par Gibbons, Ross, et Shanken (1989) comparent la performance des portefeuilles pré-composés avec celui du marché.

D'un point de vue méthodologique, les "*time series tests*" consistent à estimer le modèle suivant :

$$R_i - R_f = a_i + \beta_{i,m}(R_m - R_f) + \varepsilon_{i,t} \quad (1.4)$$

Où R_i désigne la rentabilité du portefeuille i , R_f la rentabilité sans risque. Dans les tests du CAPM, l'écart $R_i - R_f$ est souvent nommé la rentabilité excessive⁹ du portefeuille i , et $R_m - R_f$ la prime de risque du marché. $\beta_{i,m}$ dénote toujours la

9. Remarquons que, ici, la notion de rentabilité excessive se distingue catégoriquement de la définition retenue dans nos développements théoriques explicitée dans la section 1.1.1 à la page 13. La rentabilité excessive mentionnée dans les tests "*time series*" n'est soumise à aucune notion d'équilibre, elle désigne tout simplement la différence entre la rentabilité d'un portefeuille risqué et celle d'un actif sans risque.

1.2. Absence de consensus sur le modèle d'équilibre, une source de confusion

sensibilité du portefeuille par rapport aux variations de marché. $\varepsilon_{i,t}$ est un terme d'erreur souvent supposé $i.i.d.N(0, \Omega)$, $\Omega < \infty$.

Sous le CAPM, le paramètre estimé a_i ne doit pas être significativement supérieur à 0.

Utilisant les rentabilités hebdomadaires durant la période 1926 – 1966, Black, Jensen, et Scholes (1972) confirment l'équivalence entre a_i et zéro. En revanche, le CAPM est mis en question par un autre constat de ces mêmes auteurs :

- Les portefeuilles volatiles (donc ceux qui ont des $\beta_{i,m}$ plus élevés) ont souvent un $a_i < 0$,
- tandis que les portefeuilles moins risqués ont généralement un $a_i > 0$.

Ce constat, illustré par la figure 1.2, est confirmé par Fama et French (2004) avec les rentabilités mensuelles des actions américaines durant la période de 1928 à 2003. La corrélation entre a_i et $\beta_{i,m}$ constitue sans doute une menace pour la va-

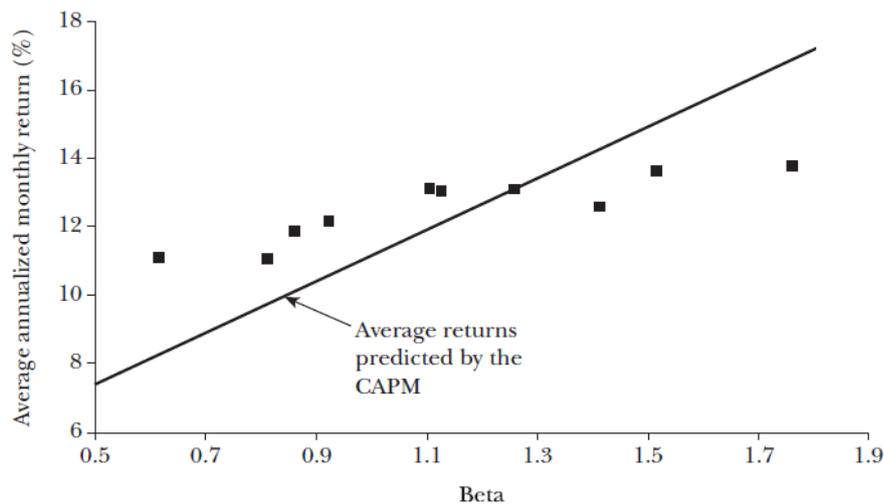


Fig. 1.2: Corrélation entre $\beta_{i,m}$ et a_i .

L'axe X représente le niveau du β , et l'axe Y la moyenne des rentabilités réalisées. La ligne droite représente les rentabilités théoriques calculées selon le CAPM.

Source : figure II Fama et French (2004), page 33.

lidité du CAPM. Toutefois, Gibbons, Ross, et Shanken (1989) ont mis en cause la robustesse d'un tel résultat en soulevant le problème des tests **ex post**. Selon ces auteurs, observer un écart **ex post** entre le CAPM et les rentabilités historiques n'implique pas toujours un échec **ex ante** du CAPM. La surperformance ou sous-performance des portefeuilles pré-composés peut être due au simple hasard, et des tests statistiques sont nécessaires pour vérifier la robustesse des observations **ex post**.

À cet effet, Gibbons, Ross, et Shanken (1989) (désormais GRS) ont introduit un test dont le déroulement est détaillé dans l'annexe C. En étudiant les rentabilités des actions américaines durant la période 01/1931 – 12/1965, ces auteurs confirment la dominance du portefeuille de marché.

Certes, en tenant compte de la nature **ex post** des observations, le test GRS présente une nette amélioration par rapport aux premiers “*time series tests*”. Cependant, comme la relation entre $\beta_{i,m}$ et a_i n'est pas observée dans le test GRS, la corrélation entre a_i et $\beta_{i,m}$ constatée par Black, Jensen, et Scholes (1972) est en réalité “masquée” mais non expliquée par ce dernier travail.

Il demeure donc probable que cette corrélation indique l'existence d'autres facteurs explicatifs dans les rentabilités d'équilibre¹⁰. Cette piste est d'ailleurs exploitée avec succès par Fama et French (1993) qui proposent un modèle d'équilibre à trois facteurs, par la suite étendu par Carhart (1997) par exemple.

III. Les modèles multi-facteurs

La recherche des facteurs explicatifs dans les rentabilités financières a commencé vers la fin des années 70. Basu (1977) montre que les entreprises avec un

10. Roll (1977) explique cette corrélation par l'observation biaisée des rentabilités du marché. Selon cet auteur, la rentabilité et le risque du marché sont des variables non observables : en théorie, le portefeuille de marché doit contenir tous les investissements financiers disponibles dans le monde entier. Or, sur le plan pratique, un tel portefeuille n'est évidemment pas observable. Par exemple, considérer les actions américaines comme les seuls titres financiers que les investisseurs puissent acheter et vendre provoque sans doute des biais d'estimation dans les tests du CAPM. À cause de ces biais, nommé “biais domestiques”, l'origine des anomalies observées demeure difficile à identifier.

Selon Stambaugh (1982), le fait d'inclure uniquement des actions dans le portefeuille de marché n'influence pas le résultat des tests du CAPM car la volatilité du vrai portefeuille de marché provient principalement du marché des actions.

Certes, comme beaucoup de tests empiriques en finance, le résultat de Stambaugh (1982) peut aussi être critiqué d'une manière ou d'une autre.

Or, même en ignorant le travail de Stambaugh, la conclusion de Roll (1977) sur irréfutabilité du CAPM paraît toujours assez catégorique. Bien que l'observation du portefeuille de marché soit soumise au biais domestique, il serait possible de rejeter le CAPM si d'autres facteurs explicatifs sont nécessaires dans le calcul des rentabilités d'équilibre. Par exemple, dans la prochaine section, il est montré que la taille d'une entreprise donnée peut influencer sa rentabilité d'équilibre (Banz 1981). Cela serait illogique de supposer qu'en introduisant les marchés internationaux dans le portefeuille de marché, le CAPM puisse prendre en compte l'effet de taille, car cela revient à supposer une corrélation entre les marchés internationaux et la taille des entreprises américaines.

Étant donnés ces faits constatés autour du CAPM, les modèles multi-facteurs sont proposés pour mieux interpréter la rentabilité d'équilibre. Ces modèles qui sont principalement fondés sur les observations empiriques feront notre sujet de discussion dans la section III.

“*price earning ratio*¹¹” (désormais PER) plus élevé engendrent, en moyenne, une rentabilité plus importante que la valeur théorique prévue selon le CAPM. Banz (1981) remarque que les petites entreprises (en termes de la capitalisation boursière) ont tendance à être plus performantes que le marché. Stattman (1980) et Fama et French (1992) soulignent l’effet du “*book to market value*¹²” (désormais B/M) sur la rentabilité espérée d’une entreprise.

Afin de prendre en considération ces faits empiriques, Fama et French (1993) ont proposé un modèle à trois facteurs qui a nettement amélioré la caractérisation des rentabilités historiques.

Selon Fama et French (1993), la rentabilité d’un actif donné est expliquée par trois facteurs : la rentabilité du marché, la taille de l’entreprise et le ratio “*book to market*”.

Plus précisément, la rentabilité d’un actif i à l’instant t peut être calculée par l’équation suivante :

$$E_{it} - R_{ft} = \beta_{im}(E_{mt} - R_{ft}) + \beta_{is}E(SMB_t) + \beta_{ih}E(HML_t) \quad (1.5)$$

Ici,

- E_{it} désigne la rentabilité espérée de l’actif i à l’instant t . R_{ft} la rentabilité sans risque au même instant.
- E_{mt} la rentabilité espérée du marché à l’instant t et β_{im} la sensibilité de l’actif i aux variations du marché.
- le facteur “taille”, $E(SMB_t)$ (SMB signifie “*small minus big*”) qui désigne l’écart de rentabilité entre les petites et grandes entreprises. β_{is} est la sensibilité de l’actif i à cette différence.
- le facteur “croissance”, $E(HML_t)$ (HML signifie “*high minus low*”), qui désigne l’écart de rentabilité entre les entreprises présentant un ratio “*book to market*” élevé à celles des sociétés dont ce ratio est faible. β_{ih} est la sensibilité de l’actif i vis-à-vis de cette différence.

Dans une étude de comparaison, Fama et French (1996) soulève que l’écart entre le CAPM et les données historiques est trois fois plus important que celui du modèle à trois facteurs¹³, et ce résultat est particulièrement net quand il s’agit des portefeuilles dotés d’un B/M élevé.

La qualité empirique du modèle à trois facteurs ne se limite pas aux comparaisons avec le CAPM. Dans ce même travail, Fama et French ont également montré

11. Le “*price earning ratio*”, est le rapport entre le cours boursier d’une entreprise donnée et son bénéfice par action.

12. Le “*book to market value*” d’une entreprise donnée est le ratio entre la valeur comptable de ses actions (les capitaux propres) et leur valeur boursière, $B/M = \frac{\text{capitaux propres par action}}{\text{prix de marché}}$.

13. L’écart mesuré par la moyenne des distances entre les rentabilités observées et leurs valeurs estimées avec les modèles respectifs.

la capacité de leur modèle à absorber “l’effet du ratio prix sur bénéfice, trésorerie sur prix, la croissance des ventes, et la rentabilité historique à long terme (Fama et French 1996, p.76).”¹⁴

Néanmoins, le modèle à trois facteurs n’est pas un modèle parfait. Certains faits empiriques échappent toujours à ce modèle. Par exemple, Jegadeesh et Titman (1993) soupçonnent une auto-corrélation positive entre les rentabilités à court terme¹⁵. En effet, ces auteurs ont découvert qu’un portefeuille ayant une rentabilité plus élevée que la moyenne du marché a tendance à préserver cette performance pendant la période suivante. Ce phénomène, connu sous le nom de l’effet “*momentum*” (Goetzmann et Massa 2002, Korajczyk et Sadka 2004), n’est pas expliqué par le modèle à trois facteurs. Selon Fama et French, les portefeuilles aux rentabilités passées élevées deviennent moins sensibles au facteur HML, ce qui impliquerait une rentabilité espérée plus basse dans l’avenir. On devrait donc observer le contraire de l’effet “*momentum*”.

Comme Fama et French (1996) l’ont précisé dans leur conclusion,

The three-factor risk-return relation is, however, just a model. It surely does not explain expected returns on all securities and portfolios.
(Fama et French 1996, 85)

Fondé sur des données mais non un raisonnement théorique, la robustesse du modèle à trois facteurs reste controversée :

- Sans support théorique, le modèle de Fama et French (1993) n’est donc pas à l’abri du risque de data-snooping (Lo et MacKinlay 1990).
- Selon Kothari, Shanken, et Sloan (1995), les deux facteurs supplémentaires (SMB et HML) sont dus au biais de sélection : les entreprises utilisées dans les tests de Fama-French doivent être présentes dans les deux bases de données sélectives : *CRSP* pour les prix observés et *Compustat* pour les caractéristiques de l’entreprise. Le modèle n’a donc pas été testé sur l’ensemble des actions disponibles. Ce biais de sélection pourrait être l’explication des résultats constatés.

Sans fondement théorique, l’interprétation du modèle multi-facteur présente également des difficultés conceptuelles. Est-il rationnel pour les investisseurs de valoriser d’autres facteurs explicatifs que le risque du marché ? La réponse à cette question dépend évidemment de la rationalité retenue.

Pour Lakonishok, Shleifer, et Vishny (1994), l’écart entre le CAPM et les données historiques est une preuve d’irrationalité : puisque les agents rationnels maximisent leur rentabilité espérée en minimisant l’écart-type associé, les facteurs comme la taille d’une entreprise ne doivent pas être pris en compte par le modèle

14. Texte original en anglais : “*patterns on average return produced by forming portfolios on E/P ($\frac{\text{earnings}}{\text{price}}$), C/P ($\frac{\text{cashflow}}{\text{price}}$), sales growth, and long term past return*”.

15. i.e. les rentabilités mesurées avec un intervalle de 3 à 24 mois.

d'équilibre.

Si, comme Fama et French, on considère la taille d'une entreprise et le ratio "book to market" comme des indicateurs de risque, alors il serait rationnel d'observer ces derniers facteurs dans un modèle d'équilibre. Dans ce cas de figure, on peut rejeter le CAPM tout en validant l'efficience informationnelle.

Conclusion

L'objectif de cette section était de présenter les principaux modèles disponibles pour calculer la rentabilité d'équilibre r_{t+1}^* . Après une investigation sur leur validité, on constate qu'aucun des deux modèles étudiés n'est en mesure de caractériser parfaitement des données empiriques.

Il est important de préciser ici que depuis le travail de Fama et French (1993), de nombreuses pistes ont été exploitées pour améliorer l'efficacité du modèle à trois facteurs. Comme Brandouy, Briec, Kerstens, et de Woestyne (2010) l'ont résumé, trois directions sont envisageables pour réduire l'écart entre le modèle d'équilibre et des données historiques :

1. L'usage de la distribution non normale des rentabilités financières. Si les rentabilités sont distribuées selon une loi non normale dont la caractérisation nécessitent plus d'indicateurs que l'espérance et l'écart-type, alors les moments supérieurs de cette distribution pourraient être les nouveaux facteurs de risque qui ont un effet sur la rentabilité d'équilibre (Harvey et Siddique 2000, Chung et Schill 2006).
2. La prise en compte de la nature "ex post" des méthodes empiriques. Une telle logique rejoint évidemment la logique du test GRS (Gibbons, Ross, et Shanken 1989, Brandouy, Briec, Kerstens, et de Woestyne 2010).
3. L'ajout de nouveaux facteurs explicatifs. Par exemple, Ross (1976) propose le modèle APT ("*arbitrage pricing theory*"), dont le nombre de facteurs retenus peut s'adapter à la base de données étudiée. Carhart (1997), Brav, Geczy, et Gompers (2000) et Tai (2003) utilisent un modèle à 4 facteurs (rentabilité de marché, taille, "book-to-market", et "winners minus losers") dans leur estimation de la rentabilité d'équilibre, afin de prendre en considération l'effet "*momentum*", une des anomalies la plus reconnue du modèle de Fama-French.

Ces modèles multi-facteurs ne sont pas développées en détail, car de notre point de vue, ils s'inscrivent à la même logique du travail de Fama et French (1993), dans la mesure où les mêmes critiques peuvent être formulées :

- ces modèles sont tous fondés sur les faits empiriques, leur interprétation théorique reste toujours discutable.

1.2. Absence de consensus sur le modèle d'équilibre, une source de confusion

- En adoptant un critère purement empirique, l'appréciation du "bon" modèle devient arbitraire. Un modèle explique les données historiques avec un R^2 à 70%, est-ce suffisamment pour le considérer comme "le modèle d'équilibre ultime" ? Sans éclaircissement théorique, la recherche d'anomalies et de nouveaux facteurs explicatifs semble être une poursuite sans fin qui rendrait discutable l'usage de tout modèle spécifique.

Les tests d'efficience qui nécessitent le calcul des rentabilités exigées doivent donc toujours supposer la validité du modèle employé. Ce problème d'hypothèse jointe a été amplement souligné dans cette section.

Le problème de l'hypothèse jointe relève également le lien théorique entre l'efficience informationnelle et le CAPM. Le CAPM est un des modèles d'équilibre sous l'efficience. Il devient le seul modèle rationnel uniquement si l'on considère l'écart-type des rentabilités comme le seul indicateur des risques financiers. Sans cette hypothèse contraignante, l'efficience informationnelle peut rester valable sans que le CAPM le soit. Cette conclusion suggère une implication importante sur les limites théoriques des fonds d'investissement qui sera développée dans la prochaine section.

1.2.2 Limites théoriques des fonds d'investissement : dans quelle mesure les marchés financiers sont-ils imbattables ?

Après avoir distingué le CAPM de l'efficience informationnelle, dans cette section nous mettons en évidence l'implication d'une telle distinction sur les performances théoriques des fonds alternatifs. Cette discussion représente un intérêt particulier sur la gestion des fonds d'investissement, notamment sur l'interprétation de leur performance à long terme.

Dans la littérature en finance, les performances des fonds sont souvent étudiées d'un point de vue empirique. Les rentabilités réalisées par des fonds d'investissement, comme proposées dans la base de données "*Micropal*", sont les objets d'études fréquemment recensés.

Par exemple, en examinant 143 fonds d'investissement pendant la période de 1965 à 1985, Ippolito (1989) constate une surperformance significative des fonds par rapport à l'indice SP500. Et ce, même après la déduction des frais de transaction. En examinant la même base de données, Elton, Gruber, Das, et Hlavka (1993) remarque que le constat d'Ippolito est expliqué par l'effet de taille. En observant une persistance dans la performance des fonds mutuels¹⁶, Grinblatt et

16. Dans la mesure où les meilleurs fonds pendant une période donnée ont tendance à conser-

Titman (1992) défendent la capacité des fonds à engendrer des gains anormaux. Chevalier et Ellison (1999) signalent que les gestionnaires de fonds diplômés des universités prestigieuses engendrent systématiquement des rentabilités excessives. En analysant les performances des fonds durant la période 1975-1996, Wermers (2000) confirme la valeur ajoutée des gérants de portefeuille actifs après les frais de gestion. Un résultat similaire est obtenu par Liang (2001) pour une période d'analyse plus récente (1990-1999). En appliquant les tests de "bootstrap", Kosowski, Timmermann, Wermers, et White (2006) valident la capacité des fonds mutuels à sélectionner les actifs rentables pendant la période de 1975 à 2002.

Par ailleurs, non seulement les travaux empiriques, surtout ceux qui ont apparu après les années 90, ont tendance à valoriser le travail des fonds actifs, certaines légendes de "Wall Street" font aussi propager la conviction en leurs stratégies d'investissement.

Pour la période 1998-2008, le fonds alternatif "Medaillon" dirigé par James H. Simons, un ancien professeur en mathématiques, a enregistré une rentabilité moyenne de 35,6% après frais de gestion, tandis que l'indice SP500, durant cette même période, n'a augmenté qu'en moyenne de 9,2% par an. En 2008, durant la crise immobilière, ce fonds a emporté un rendement net de 80%.

Face à ces performances "extraordinaires", l'efficience informationnelle est souvent défendue par deux arguments théoriques :

- *phénomène de chance* : si les rendements des fonds sont distribués de façon aléatoire, alors, en observant un grand nombre de fonds, il est possible de constater certaines performances "anormales" sans remettre en question l'efficience informationnelle, car celles-ci sont toujours accompagnées d'un grand nombre de fonds "normaux" dont le rendement est bien inférieur à celui de marché et la moyenne du métier reste plus que raisonnable.
- *ajustement du risque* : selon le CAPM, l'écart de performance entre un fonds alternatif et un indice de marché peut provenir de deux sources différentes : la rentabilité anormale (le " α "), et la rentabilité provoquée par le risque (le " β ").

Comme les stratégies pratiquées par les fonds alternatifs sont couvertes par le secret professionnel, le niveau de risque subi par ceux-ci est donc difficile à estimer. De ce fait, distinguer le " α " du " β " reste une tâche extrêmement délicate.

Dans cette section, nous mettons en avant une autre explication à la performance de fonds tel que "Medaillon", en distinguant les stratégies de "market timing" de celles dites de "stock picking". Bien que le modèle parfait reste à trouver, en acceptant la remise en cause du CAPM, on pourrait espérer un α positif en pratiquant des stratégies de "stock picking".

ver cette performance pendant la période suivante.

“*Sur un marché efficient, on ne peut pas battre le marché*”. Cette implication de l’efficience informationnelle semble partagée par la majorité des chercheurs en finance et a fait l’objet de nombreux tests empiriques. “Battre le marché” est défini de manière générale comme l’absence de rentabilités excessives à long terme. Deux significations différentes sont souvent associées à ce terme¹⁷ :

- “Battre le marché” en entrant et sortant du marché aux meilleurs moments (“*market timing*”). Pour tester cette possibilité, on compare la performance de stratégies de type “*market timing*” avec celle de “*buy and hold*”, i.e., une stratégie qui consiste à détenir un portefeuille donné sans jamais sortir du marché.
- “Battre le marché” en choisissant les titres les plus rentables (“*stock picking*”). Dans ce cas de figure, la comparaison se fait entre la performance des stratégies “*stock picking*” et celle du portefeuille de marché.

Dans la majorité des tests empiriques, battre le marché dans ces deux sens constitue une preuve à l’inefficience des marchés financiers.

Cependant, ces deux significations pour “battre le marché” ne partagent pas les mêmes fondements théoriques. L’impossibilité du “*marketing timing*” est une implication directe de l’efficience informationnelle, alors que celle du “*stock picking*” est une conséquence du CAPM. Comme l’efficience peut être vérifiée sans que le CAPM le soit, cette dernière contrainte est plus “stricte” que la précédente.

Cette distinction est particulièrement intéressante étant donné que la validité du CAPM est déjà mise en question par de nombreux faits empiriques :

- Sur le plan pratique, cette distinction remet en question les méthodes traditionnellement utilisées pour évaluer les performances de fonds, notamment les ratios de Sharpe et Treynor, ou encore l’indice Alpha de Jensen. Un gestionnaire de fonds qui surperforme le portefeuille de marché ne met qu’en cause le CAPM mais non l’efficience informationnelle. Le vrai défi théorique se trouve dans la possibilité de “*market timing*”.
- Sur le plan théorique, notre analyse permet d’expliquer, en préservant l’efficience informationnelle, la performance de certains fonds d’investissement qui semblent dépasser la rentabilité du marché sur le long terme.

De plus, l’impossibilité de “*marketing timing*” est pour le moment, la seule implication de l’efficience informationnelle directement testable avec les données empiriques. Comprendre cela permet de mieux interpréter les résultats empiriques observés sur le marché.

17. Remarquons que le premier article qui propose une telle séparation fut Chang et Lewellen (1984).

I. "Market timing"

Comme nous l'avons évoqué précédemment, la rentabilité d'un actif est composée de deux éléments distincts : la rentabilité d'équilibre (r_t^*) et la rentabilité excessive ($\varepsilon_{t+1} = r_{t+1} - r_t^*$) qui vérifie une espérance conditionnelle nulle.

Selon Samuelson (1965), r_{t+1}^* ne peut être négative, car les investisseurs n'ont aucune raison de demander une récompense négative pour la détention d'un actif financier. Cette condition ($r_{t+1}^* \geq 0$) est non seulement soutenue par la plupart des théoriciens en finance, mais est aussi compatible avec les analyses des investisseurs professionnels (Arnott et Bernstein 2002).

Remarquons que cette non-négativité est remise en question par Fama, et ce à plusieurs reprises. Fama(1970) justifie la possibilité d'observer $r_{t+1}^* < 0$ avec un argument fondé sur le CAPM :

"Negative equilibrium expected returns for some securities are quite possible. For example, in the Sharpe (1964) - Lintner (1965) model, the equilibrium expected returns on a security depends on the extent to which the dispersion in the security's return distribution is related to dispersion in the returns on all other securities. A security whose returns on average move opposite to the general market is particularly valuable in reducing dispersion of portfolio returns, and so its equilibrium expected return may well be negative (Fama 1970, p.289)."

Selon Fama, même en anticipant une baisse importante d'un tel titre, les agents risque-averses ne voudraient pas le vendre pour diversifier le risque du portefeuille. Or, la théorie du portefeuille est fondée sur la diversification des risques spécifiques, non celle de risque systémique. La contribution d'un titre i à la diversification du risque dépend de sa corrélation avec les autres titres individuels. La covariance entre i et le marché est un indicateur du risque systémique qui ne mesure pas toujours son effet de diversification. Autrement dit, les titres diversifiant au mieux un portefeuille donné n'engendrent pas toujours des rentabilités négatives. Les investisseurs n'ont donc pas de raison valable de garder les titres perdants dans leur portefeuille.

La non-négativité de r_t^* sera donc retenue dans les analyses des performances de fonds.

Sous cette condition, la limite des stratégies du type "market timing" est défendue dans l'article de Fama (1970). En effet, sous la condition $r_t^* \geq 0$, Fama affirme que

"the set of 'one security and cash' mechanical trading rules by which we mean systems that concentrate on individual securities and

that define the conditions under which the investor would hold a given security, sell it short, or simply hold cash at any time t , cannot have greater expected profits than a policy of always buying-and-holding the security during the future period in question. (Fama 1970, p.386)

Dans le système de “trading” défini par Fama, la rentabilité accumulée d’un investisseur est la somme de toutes les rentabilités réalisées pendant les périodes où il détient l’actif. Comme les rentabilités excessives sont imprévisibles, elles vont s’annuler dans le temps, la somme des rentabilités d’équilibre est le seul élément qui détermine la rentabilité à long terme.

Étant donné $r_t^* \geq 0$ pour tout instant t , la meilleure stratégie est de garder le titre à long terme sans jamais sortir du marché, i.e., la stratégie “*buy and hold*”.

II. “Stock picking”

Si l’impossibilité du “*market timing*” provient de la loi de l’offre et de la demande, l’inefficacité de “*stock picking*” est une implication du CAPM, un des modèles d’équilibre sous l’efficience informationnelle. Si, comme beaucoup d’études empiriques l’ont soulevé, le marché est efficient avec un autre modèle d’équilibre que le CAPM (Fama et French 1993, Elton, Gruber, Das, et Hlavka 1993, Carhart 1997), alors l’efficience informationnelle peut tout à fait être vérifiée sans la supériorité du portefeuille de marché.

Fama parle de “*multifactor efficiency*(Fama et French 1996, p.444)” pour se démarquer du “*mean-variance efficiency* (Fama et French 1996, p.444)” au sens du CAPM. Selon l’auteur, si les investisseurs n’utilisent pas l’écart-type comme le seul indicateur du risque, et l’espérance comme l’unique mesure de la rentabilité, la dominance du portefeuille de marché ne serait plus fondée. Il devient donc possible des composer des portefeuilles actifs engendrant un ratio espérance/écart-type supérieur à celui du marché.

Comme l’imprévisibilité des rentabilités excessives est soumise à moins d’hypothèses théoriques que le CAPM, il doit être moins difficile de battre le marché dans le sens du “*stock picking*” que dans le sens du “*market timing*”.

Cette conclusion est aussi bien confirmée par les études quantitatives que qualitatives :

- Daniel (1997) montre que même en prenant comme benchmark le modèle à 4 facteurs de Carhart (1997), les fonds d’investissements engendrent des rentabilités anormales avec les techniques de “stock picking”.
- Dans un travail qualitatif publié par Shiller (2005), cet auteur remarque :

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

“When he (Shiller) asked his respondents to finish the phrase : 'Trying to time the market to get out before it goes down and to get in before it goes up is ...', only 11 percent chose : 'a smart thing to try to do ; I can reasonably expect to be a success at it.' Yet when he asked them to complete this phrase : 'Trying to pick individual stocks, trying to predict, for example, if and when Ford Motor stock will go up or IBM stock will go up is ...,' 40 percent of the respondents chose 'a smart thing to do, I can reasonably expect to be a success at it.' (Bernstein 2007, p.75-76)”

La dissociation entre le CAPM et l'efficacité informationnelle a un effet indéniable sur la notion de “battre le marché”. L'impossibilité de “*marketing timing*” peut être valable sans que les stratégies “*stock picking*” soient inefficaces. Il est moins difficile de battre le portefeuille de marché que la stratégie “*buy and hold*”.

Sur le plan théorique, cette distinction contribue au débat en cours concernant l'efficacité informationnelle. Étant donné le problème de l'hypothèse jointe, il semblerait que la seule façon de rejeter cette dernière théorie serait de montrer la réussite d'une stratégie “*marketing timing*” à long terme. Ce critère sera utilisé dans la prochaine section pour analyser les implications théoriques des anomalies du marché.

1.3 Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

Depuis sa première apparition dans les années 70, l'efficacité informationnelle semble être remise en question par de nombreux faits empiriques, nommés “anomalies du marchés”. Ces anomalies seront analysées dans cette section en soulignant leurs impacts sur la prévisibilité des rentabilités excessives.

Comme nous l'avons évoqué précédemment, 3 nuances théoriques sont à distinguer pour bien interpréter les observations empiriques :

1. le CAPM et l'efficacité, autrement dit, la distinction entre la réussite des stratégies de “*marketing timing*” et de “*stock picking*”.
2. le rejet de l'hypothèse jointe et de l'efficacité informationnelle.
3. les significations statistique et économique. Comparées avec le marché libre décrit par la théorie, les bourses réelles sont soumises à de nombreuses contraintes pratiques. Étant donnée la présence des coûts de transaction, la durée limitée d'un horizon d'investissement, etc, “battre le marché” en théorie n'implique pas toujours la faisabilité pratique. Les marchés financiers,

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

qui ne sont pas les marchés parfaits, peuvent être qualifiés "d'efficent", si les rentabilités excessives sont irréalisables en pratique.

Ces 3 principes seront appliqués dans nos analyses des anomalies du marché. La conclusion dressée à l'issue de ces analyses va en faveur de la théorie de l'efficence dans la mesure où aucun travail ne montre, à notre connaissance, l'existence d'une stratégie de "*marketing timing*" qui peut battre la "*buy and hold*" en pratique.

1.3.1 Effet calendrier

L'effet calendrier est défini par la sensibilité des rentabilités espérées au calendrier civil.

- Rozeff et Kinney (1976) et Haugen et Lakonishok (1988) ont remarqué un possible "effet janvier", qui se traduit par une rentabilité plus élevée en janvier que pendant les autres mois de l'année.
- Cross (1973) et French (1980) ont quant à eux signalé un "effet week-end" défini par une rentabilité moyenne plus élevée le vendredi que le lundi.

Certains auteurs ont constaté d'autres relations entre le calendrier et les rentabilités financières, notamment un possible "effet de vacance", "effet de fin de mois", ou bien encore "effet d'intra-day" (Thaler 1987, p.173). La plupart de ces constats restant très isolés et sans vérification ultérieure, nous nous focaliserons sur l'effet janvier et l'effet week-end en essayant de mettre en évidence leur origine et leurs impacts sur la théorie de l'efficence.

Effet janvier La discussion sur "l'effet janvier" a commencé par le travail de Rozeff et Kinney (1976). En étudiant un indice équi-pondéré, qui contient toutes les actions cotées sur le NYSE pendant la période 1904-1974, les auteurs ont reporté une rentabilité moyenne de 3.84% pour le mois janvier, contre 0.42% durant le reste de l'année. Or, les études ultérieures (Keim 1983, Reinganum 1983, Roll 1983) ont montré la relation négative entre la taille d'une entreprise et sa rentabilité en janvier. Depuis ces derniers travaux, l'effet janvier est considéré comme un phénomène réservé aux entreprises de petite taille.

Selon Ritter (1988), ce phénomène s'explique par les comportements des investisseurs individuels au tournant de l'année. En effet, à la fin de chaque année, les investisseurs individuels cherchent à vendre les actifs en perte dans leur portefeuille pour réaliser une moins-value fiscale. Ceux-ci vont attendre janvier pour réinvestir le capital ainsi libéré de sorte à intégrer dans le même temps les primes de fin d'année qu'ils obtiennent de leur entreprise. Comme les petites entreprises sont principalement détenues par les investisseurs individuels, l'effet janvier engendré par ce comportement ne concerne que les entreprises de petite taille.

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

Cette explication qui combine la fiscalité et les comportements des investisseurs individuels a été confirmée par des travaux plus récents. Parmi ceux-ci, D'Mello, Ferris, et Hwang (2003) ont par exemple observé une grande pression de vente sur les actifs perdants à la fin de l'année, et une attente volontaire de la part des investisseurs individuels pour renouveler leurs achats.

Sur le plan théorique, "l'effet janvier" est une anomalie dans le sens du CAPM mais non une preuve d'inefficience :

1. il peut être expliqué par un comportement rationnel des investisseurs : réaliser les moins-values fiscales pour réduire ses impôts sur les revenus. Bien que le phénomène ne soit pris en compte par aucun modèle d'équilibre actuellement disponible, il n'a rien d'un facteur psychologique ou d'un processus autoréférentiel.

En revanche, "l'effet janvier" peut éventuellement être considéré comme une anomalie du CAPM, car il soutient le fait que l'espérance et l'écart-type des rentabilités ne sont pas les seuls facteurs pris en compte dans les décisions d'investissement.

2. l'excès de rentabilité observé en janvier n'est pas suffisamment important pour engendrer les gains anormaux, surtout en présence de coûts de transactions. Étant donné le développement des marchés financiers et surtout l'amélioration rapide des conditions de change, le soi-disant "effet janvier" semble avoir disparu peu de temps après sa découverte (Malkiel 2003, p.60)¹⁸.

Ce point de vue est confirmé par le travail de Szakmary et Kiefer (2004), qui montre la disparition de "l'effet janvier" après l'introduction des contrats "future SP Midcap" et "Russell 2000" en 1993. En effet, ces deux produits dérivés ont été conçus dans l'objectif de faciliter l'arbitrage sur les titres des petites et moyennes entreprises¹⁹. La disparition de cet effet suite à un arbitrage plus accessible est sans aucun doute un argument en faveur de la rationalité.

3. comme l'effet janvier est un phénomène réservé aux petites entreprises, on peut donc le saisir avec le modèle à trois facteurs (Fama et French 1993). En le considérant comme un modèle rationnel, l'augmentation de rentabilité en janvier devient une implication naturelle du modèle d'équilibre, mais non une preuve d'inefficience.

Effet week-end "Effet week-end", initialement découvert par Cross (1973) et confirmé par French (1980), Gibbons et Hess (1981), Keim et Stambaugh (1984)

18. Texte original en anglais : "the so called "January effect", ..., seems to have disappeared soon after it was discovered (Malkiel 2003, p.60)."

19. L'avantage d'exploiter "l'effet janvier" sur le marché à terme a été démontré par Clark et Ziemba (1987).

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

ou bien encore Jaffe et Westerfield (1985) est défini par le fait que les rentabilités financières sont en moyenne plus élevées le vendredi que le lundi. Contrairement à "l'effet janvier", à notre connaissance, il n'existe pas d'explication unanimement acceptée à ce phénomène empirique. Selon les défenseurs de la finance comportementale, les investisseurs sont prêts à offrir un meilleur prix le vendredi, grâce à leur disposition d'esprit plus positive avant le week-end (Thaler 1987).

Dans l'ignorance de son origine, la relation entre l'effet "week-end" et la rationalité est difficile à déterminer. Néanmoins, comme "la différence de rentabilité (entre vendredi et lundi) est faible (Simon et Joffre 1997, p.429)" et que le "jour noir" de la semaine peut varier selon le marché étudié (Martikainen et Puttonen 1996), en présence de coûts de transactions, on pourrait conclure sur l'impossibilité de "*market timing*" sans prendre beaucoup de risque. L'effet "week-end" ne semble donc pas être une anomalie réfutant l'efficience informationnelle.

1.3.2 Effet "*momentum*"

Observés à une fréquence de 3 à 24 mois, les portefeuilles engendrant une rentabilité supérieure (inférieure) à la moyenne du marché ont tendance à conserver cette performance pendant la période suivante. Cette anomalie, initialement constatée par Jegadeesh et Titman (1993) durant la période de 1965 à 1989, est connue sous le nom de l'effet "*momentum*".

Ce résultat confirmé par Jegadeesh (2001) avec des données des années 90, et Rouwenhorst (1998) en testant les marchés internationaux semble indiquer l'existence d'une stratégie engendrant des rentabilités excessives. Depuis que Korajczyk et Sadka (2004) ont démontré l'efficacité des stratégies "*buy the winner and sell the loser*" après coûts de transaction, l'effet "*momentum*" est devenue l'anomalie la plus citée pour contester la théorie de l'efficience.

Par exemple, Cooper, Gutierrez, et Hameed (2004) et Hong, Lim, et Stein (2000) ont découvert que l'effet "*momentum*" dépend de la taille des entreprises et du suivi des analystes financiers. Ces auteurs ont ainsi conclu à une transmission progressive des informations pertinentes dans les cours, qui entraîne une sous-réaction ou une sur-réaction des investisseurs.

Cependant, les facteurs psychologiques ne sont pas les seules explications de l'effet "*momentum*". Il se pourrait que la taille d'une entreprise et le suivi des analystes financiers soient des indicateurs de risque choisis par les investisseurs rationnels pour calculer leur rentabilité exigée. En raison du problème de l'hypothèse jointe, le partage entre l'efficience informationnelle et la finance comporte-

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

mentale reste difficilement à établir en ce sens.

En revanche, il est toujours possible de réfuter l'efficience en identifiant une stratégie du type "*market timing*" qui battrait la stratégie "*buy and hold*".

Il est donc important de remarquer ici que les travaux empiriques qui confirment l'effet "*momentum*" sont plus dans la logique du "*stock picking*" que du "*market timing*". Plus précisément, pour tester cet effet, Jegadeesh et Titman (1993) divisent leur période d'observations (1965-1989) en un certain nombre de sous-périodes qui durent entre 3 et 12 mois. La stratégie testée consiste à acheter, à chaque sous-période, les 10% des actifs les *plus* rentables pendant la période précédente et à vendre les 10% les *moins* rémunérateurs. Cette stratégie semble pouvoir engendrer une rentabilité plus élevée que le portefeuille de marché.

Selon Lo et MacKinlay (1990), le succès des "*buying winner and selling loser*" peut provenir de trois sources différentes :

1. Certains titres ont une rentabilité inconditionnelle supérieure aux autres. Cette hypothèse initialement posée par Conrad et Kaul (1998) est écartée par Jegadeesh et Titman (2002), car selon ces derniers auteurs, les "*winner*" pendant une période donnée ne le restent pas pour toujours. Au bout d'un certain temps, les anciens "*winner*" deviennent les "*losers*".
2. Les rentabilités sont positivement auto-corrélées. Cette hypothèse n'est pas valable, car Lewellen (2002) a montré que, observée à des fréquences variant de trimestrielle à biannuelle, l'auto-corrélation des séries de rentabilités est "non significativement négative", mais non positive.
3. Il n'existe pas d'auto-corrélation dans les séries des rentabilités, mais les rentabilités d'un actif donné sont négativement corrélées avec les rentabilités retardées d'un autre actif (le phénomène nommé "*negative cross-serial covariance*" dans la littérature anglophone (Chen et Hong 2002)). Cette explication semble être la seule valable à notre connaissance. Du point de vue de l'efficience, le phénomène de "*negative cross-serial covariance*" est plus liée aux stratégies de type "*stock picking*" que "*market timing*".

Pour conclure, l'effet "*momentum*" n'implique ni l'irrationalité des agents, ni la possibilité de "*market timing*", cette anomalie est donc, de notre point de vue, compatible avec l'efficience informationnelle.

1.3.3 Excès de la volatilité et retour de moyenne

L'excès de volatilité initialement découvert par LeRoy et Porter (1981) et Shiller (1981) se traduit par une volatilité trop importante du prix par rapport au dividende.

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

Le retour à la moyenne (*“mean reversion”*) est défini par une auto-corrélation négative entre les rentabilités à long terme. Ce phénomène confirmé à la fois par les observations en séries chronologiques (Fama 1991) et en instantanée (DeBondt et Richard 1985), est également une des anomalies de marché.

Ces deux anomalies sont présentées ensemble dans cette section, car selon LeRoy (1989), celles-ci partagent la même origine théorique²⁰.

En acceptant le modèle de dividende, la volatilité des prix d'un actif ne doit pas dépasser celle de sa valeur fondamentale. Cette infériorité est rejetée indépendamment par LeRoy et Porter (1981) et Shiller (1981)²¹ qui constatent un écart-type beaucoup trop élevé des prix par rapport à celui des dividendes.

Ackert et Smith (1993) ont quant à eux remarqué que la distribution des bénéfices sous forme de rachat d'actions et les primes encaissées par les investisseurs lors d'une offre d'achat publique ne sont pas prises en compte par les tests de volatilité. En réintégrant ces flux financiers, l'excès de volatilité est rejeté par ces derniers auteurs.

La relation entre ce phénomène et la rationalité des agents reste donc controversée. Pour tester la compatibilité d'une telle anomalie avec l'efficacité informationnelle, il faut vérifier son impact sur la possibilité de *“market timing”*.

Quand il s'agit de concevoir des stratégies d'investissement, le phénomène d'excès de volatilité est souvent interprété sous l'angle du “retour à la moyenne”, phénomène souvent testé de deux manières différentes :

1. tester l'auto-corrélation des rentabilités à long terme (Fama et French 1988). Les travaux dans cette direction sont souvent limités par la disponibilité des données. Avec un intervalle d'observation entre 3 et 5 ans, les séries de rentabilités historiques présentent peu d'observations, et la robustesse des tests est souvent remise en question.
2. démontrer l'efficacité des stratégies *“buying loser and selling winner”*, ou les stratégies *“contrarian”* (DeBondt et Richard 1985, Lo et MacKinlay 1990, Loughran et Ritter 1996, Boynton et Oppenheimer 2006) par rapport au portefeuille du marché.

Comme pour l'effet “momentum”, le succès des stratégies *“contrarian”* n'implique pas la possibilité de faire du *“market timing”*.

En effet, selon Zarowin (1990), l'effet *“contrarian”*, “l'effet janvier” et “l'effet

20. Pour plus de détails sur la démonstration LeRoy (1989), voire l'annexe E.

21. En effet, Shiller (1988) a répondu aux critiques méthodologiques de Kleidon (1986) concernant son travail en 1981, et a confirmé la volatilité trop importante des prix.

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

de taille” sont les mêmes phénomènes empiriques observés sous des angles de vue différents. En ajustant les deux derniers effets, le phénomène du retour de moyenne disparaît. Le retour à la moyenne s'inscrit dans une logique de “stock picking” qui ne réfuterait pas l'efficience informationnelle.

1.3.4 Prime de risque négative

Selon la loi de l'offre et de la demande, en anticipant une prime de risque négative, les investisseurs devraient vendre l'actif concerné pour réduire le prix actuel, jusqu'au moment où la rentabilité anticipée deviendrait égale à la rentabilité exigée.

La prédiction des primes de risque négatives peut être analysée comme une anomalie de marché car cela met en cause l'aversion aux risques des agents rationnels. Autrement dit, un tel constat signifie que les investisseurs “acceptent” des rentabilités négatives pour la détention des actifs risqués.

Les tests de non-négativité ont été conçus pour vérifier le signe des primes de risque. Ces études, souvent fondées sur les modèles multi-facteurs, affirment la possibilité de prévoir de façon **ex ante** des primes de risque négatives (Boudoukh, Richardson, et Smith 1993, Ostdiek 1998, Walsh 2006, Eleswarapu et Thompson 2007).

L'objectif de cette section est d'étudier ces tests de non-négativité et surtout de montrer leur compatibilité avec l'efficience informationnelle.

Boudoukh, Richardson, et Smith (1993) (désormais BRS) ont simulé différentes stratégies du type “*market timing*” avec les variations mensuelles des indices américains. En examinant les primes de risque engendrées par ces stratégies, BRS ont rejeté l'hypothèse de la non-négativité. Cette conclusion est confirmée par les études de Kothari et Shanken (1997) et Baker et Wurgler (2000) en adoptant une technique de “*bootstrapping*” pour augmenter la robustesse des tests. Walsh (2006) et Eleswarapu et Thompson (2007) affirment la négativité des primes de risque en réalisant les “*out-of-sample*” tests. Certains investisseurs professionnels (Arnott et Bernstein 2002) estiment également une prime de risque nulle, voire négative, pour les marchés boursiers après 2002.

Une stratégie qui évite de détenir les actifs quand le marché offre une prime de risque négative peut éventuellement battre la “*buy and hold*”. Cependant, battre le marché en théorie n'implique pas toujours une telle possibilité en pratique.

– Premièrement, l'existence des coûts de transactions constitue un obstacle

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

indéniable aux opérations d'arbitrage. À cause des coûts de transaction, même en anticipant une prime de risque légèrement négative, les investisseurs ne peuvent pas toujours réaliser de profits anormaux dans le sens du “*market timing*”.

- Deuxièmement, les primes de risque négatives sont souvent observées pendant les périodes de crises financières où la liquidité du marché est épuisée. Par exemple, dans le travail de Eleswarapu et Thompson (2007) les 9 primes de risque négatives ont toutes été anticipées durant la période de 04/1973 à 05/1974 qui correspond à la crise pétrolière. Pendant cette période de crise, la liquidité des marchés s'évapore massivement, les transactions deviennent ponctuelles, même en anticipant une prime de risque négative, les investisseurs ne peuvent pas toujours trouver des acheteurs pour se débarrasser de leurs titres. Profiter de ces prévisions avec une stratégie de type “*market timing*” reste une tâche difficile²².

1.3.5 Sur-réaction et sous-réaction événementielles

Une autre anomalie souvent citée par les comportementalistes est soulignée par les études d'événement : la rentabilité excessive à long terme après l'annonce d'un événement financier porte, en moyenne, le même signe que la rentabilité observée le jour de la révélation (Grinblatt, Masulis, et Titman 1984, Bernard et Thomas 1990, Lakonishok et Vermaelen 1990, Michaely, Thaler, et Womack 1995).

Par exemple, Bernard et Thomas (1990) montrent que les rentabilités excessives observées 3 ans après les annonces des bénéfices (“*earnings announcements*”) portent, en moyenne, le même signe que la rentabilité réalisée le jour de leur publication. LaPorta, Lakonishok, Shleifer, et Vishny (1997) apportent plus de précision à ce résultat en rajoutant que “*value assets*” réalisent probablement une rentabilité positive le jour de l'annonce de bénéfices, tandis que “*growth assets*” ont souvent un rendement négatif. Grinblatt, Masulis, et Titman (1984) et Ikenberry, Rankine, et Stice (1996) constatent que les rentabilités excessives après des “*stock splits*”²³ sont significativement supérieures à zéro. Michaely, Thaler, et Womack (1995) affirment que les communications sur la valeur des dividendes influencent également la rentabilité excessive à long terme. Les études

22. Le rôle de la liquidité dans le mécanisme du marché est souligné dans les analyses de Bourghelle et Hyme (2009)

23. Quand le prix d'une action donnée est trop élevé par rapport aux autres, l'entreprise sous-jacente peut augmenter le nombre total de ses actions émises, en gardant la même capitalisation boursière. Cette opération fait retomber immédiatement le prix de chaque action cotée tout en préservant la valeur de l'entreprise. Par exemple, si l'action de l'entreprise *A* dépasse 100 euros, *A* peut diviser chacune de ses actions par 2. Après la division, chaque nouvelle action de *A* sera cotée à 50 euros. Cette opération est nommée “Stock split” par les chercheurs anglo-saxons.

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

d'événements sur les rachats d'entreprises (Lakonishok et Vermaelen 1990) et les "seasoned equity offerings"²⁴ fournissent des résultats similaires.

Pour autant, ces résultats ouvrent-ils la voie à un possible "market timing" ?

D'une façon générale, les études d'événements sont organisées en 4 étapes :

1. Fixer l'événement à étudier, par exemple, l'annonce du montant du dividende.
2. Calculer, pour chaque titre observé, **la rentabilité pendant l'événement**, notée $r_i = \frac{p_{j+1} - p_{j-1}}{p_{j-1}}$, où "j" désigne la date de l'événement, et **la rentabilité après l'événement**, notée $r'_i = \frac{p_{j+230 \times 3} - p_{j+1}}{p_{j+1}}$. Ici, on suppose qu'il y a 230 jours de transactions par an.
3. Calculer, pour chaque titre observé, **la rentabilité du marché pendant l'événement**, notée $r_{m,i} = \frac{p_{m,j+1} - p_{m,j-1}}{p_{m,j-1}}$, où $p_{m,j+1}$ est la valeur d'un indice boursier le jour après l'annonce, et **la rentabilité du marché après l'événement**, notée $r'_{m,i} = \frac{p_{m,j+230 \times 3} - p_{m,j+1}}{p_{m,j+1}}$.
4. La rentabilité excessive pendant (après) l'événement est calculée par $re_i = r_i - r_{m,i}$ (resp. $re'_i = r'_i - r'_{m,i}$) pour chaque titre observé.
5. La différence entre la moyenne de re'_i et 0 est testée pour fixer le signe des rentabilités excessives.

Fama (1998) conteste la signification de ces tests en soulignant la différence entre les rentabilités à long terme et à court terme. En effet, si la moyenne des r'_i réalisée pendant la première année après l'annonce d'un événement est, *par hasard*, 1% supérieure à la rentabilité du marché, cet écart pourrait être non significatif selon les tests statistiques.

Or, en supposant que la rentabilité du marché est de 6% par an, la rentabilité du marché durant les 3 ans après l'événement serait de $(1 + 6\%)^3 - 1 = 19,10\%$, tandis que la rentabilité moyenne des r'_i pendant la même période serait de $(1 + 7\%)(1 + 6\%)^2 - 1 = 20,23\%$. Si la différence initiale de 1% n'est pas significativement différente de zéro, l'écart entre 19,10% et 20,23% l'est probablement. Ainsi, en remplaçant les rentabilités à 1 an par celles à 3 ans, on augmente artificiellement la signification des rentabilités excessives. Fama conteste donc l'usage des rentabilités à long terme dans les études d'événements.

Même sans remettre en cause la signification des rentabilités excessives, on remarque que les études d'événements sont plus dans une logique de "stock picking"

24. Ou "secondary equity offering", ce terme désigne l'issue des nouvelles actions concernant les entreprises déjà publiquement cotées sur les marchés boursiers. Cette opération est le contraire de l'IPO ("Initial Public Offering") qui permet aux entreprises non cotées d'ouvrir, pour la première fois, leur capital aux investisseurs financiers.

1.3. Certaines anomalies du marché semblent contredire l'imprévisibilité des rentabilités excessives

que de “*market timing*”. Exploiter les annonces publiques avec les stratégies du type “*market timing*” reste un travail difficilement réalisable.

La stratégie “*buy and hold*” est battable, s'il est possible de prévoir les rentabilités négatives. Donc, si la publication des mauvaises nouvelles²⁵ engendre une rentabilité négative pendant les 3 ans après l'annonce, on doit pouvoir profiter du phénomène en évitant de tenir les actifs durant ces périodes. Or, les résultats des études d'événements n'impliquent pas une telle possibilité car :

1. La négativité des *rentabilités excessives anticipées* n'implique pas celle des *rentabilités réalisées*. La plupart des études d'événements utilisent la période d'observation 1976-1988 où le marché boursier américain a connu une croissance annuelle de 10% en moyenne.²⁶ À condition que sa valeur absolue ne dépasse pas 10%, la rentabilité excessive, définie dans ces études, par la différence entre la rentabilité réalisée et la rentabilité de marché, peut être négative sans que la rentabilité réalisée le soit. Donc, prévoir une **rentabilité excessive** inférieure à 0, n'implique pas une baisse de prix dans le futur. Vendre les titres après les annonces des mauvaises nouvelles ne peut pas toujours permettre de battre la stratégie “*buy and hold*”.
2. Comme Elton (1999) l'a souligné à juste titre, la moyenne des rentabilités réalisées n'est pas toujours une bonne estimation de la rentabilité espérée. Les valeurs extrêmes sont parfois responsables des conclusions obtenues.

En suivant cette logique, la négativité des rentabilités excessives peut être due à un petit nombre d'annonces publiques qui sont effectivement les premiers indicateurs d'une détresse économique ou financière de l'entreprise concernée. Les rentabilités excessives après ces signes de détresse sont considérablement inférieures à 0 car les entreprises qui rentrent dans une période difficile continueront à émettre des mauvaises nouvelles pendant un certain temps.

Par conséquent, pour exploiter le résultat des études événementielles avec une stratégie de “*market timing*”, il faudrait distinguer les vrais signes de détresse et les annonces publiques sans conséquence significative. Cela est en pratique difficilement réalisable. Une stratégie de “*market timing*” doit donc avoir un horizon d'investissement extrêmement long pour “tomber” sur suffisamment de signes de détresse et engendrer une différence de gains significative par rapport à la stratégie “*buy and hold*”. Et si la proportion des signes de détresse est peu élevée sur l'ensemble des informations négatives, le “*market timing*” peut devenir impossible en présence des coûts de transaction.

Dans cette section, nous avons étudié les principales anomalies de marché dans

25. C'est-à-dire que les nouvelles qui engendrent une rentabilité négative le jour de l'annonce.

26. Le taux de croissance est calculé avec les données historiques de l'indice Dow-Jones, pendant la période 02/01/1976-03/01/1989.

la littérature en finance. Ces anomalies sont souvent citées par les comportementalistes pour réfuter l'efficacité informationnelle. Or, il est argumenté ici que les anomalies de marché sont souvent liées aux limites du CAPM. Aucune de ces anomalies n'implique la réussite des stratégies de "*market timing*".

Cette argumentation fondée sur la loi de l'offre et de la demande, est en ligne avec notre discussion sur la validité des modèles d'équilibre. Comme il a été montré dans la section précédente, le CAPM qui fonctionne dans un monde "espérance - écart-type", présente des défauts empiriques incontestables. La notion d'équilibre multi-facteurs, dans le sens de Fama et French (1996), semble mieux adaptée aux faits empiriques. L'existence d'anomalies contre le CAPM reste, par conséquent, une observation cohérente avec les sections précédentes.

Or, faute d'un modèle d'équilibre universel, la relation entre ces anomalies et la rationalité des agents s'avère difficile à interpréter. En conséquence, l'efficacité informationnelle ne peut être rejetée qu'en démontrant la possibilité de "*market timing*". Puisque aucune anomalie étudiée ne semble impliquer une telle possibilité, la théorie de l'efficacité demeure, de notre point de vue, compatible avec ces anomalies du CAPM.

Conclusion

L'objectif de ce chapitre était de mettre en évidence la manière dont les cours boursiers évoluent d'un point de vue théorique. Pour effectuer une telle analyse, il est introduit, comme cadre général explicatif, la loi de l'offre et de la demande, principe autour duquel sont organisées la plupart des places boursières.

À l'aide de ce cadre, il a été étudié trois grands courants théoriques qui s'interrogent sur les comportements des agents économiques : la théorie de l'efficacité, la finance comportementale, et la finance conventionnaliste. Notre conclusion à l'issue de cette analyse est la suivante : l'absence de rentabilités excessives à long terme, souvent considérée comme une implication directe de l'efficacité, est en réalité compatible avec les trois théories concurrentes.

Or, si la performance des stratégies de type "*market timing*" ne permet plus de distinguer les différentes théories financières, comment peut-on réfuter l'efficacité informationnelle? Étant donné le problème de l'hypothèse jointe, le seul point à même de fragiliser l'hypothèse d'efficacité est le succès des stratégies du type "*market timing*".

La distinction entre les stratégies de “*market timing*” et de “*stock picking*” est un autre point essentiel dans la compréhension de l’efficience informationnelle. Les techniques de “*stock picking*” sont souvent utilisées pour battre la performance de marché. Cependant, la supériorité du portefeuille de marché est une implication du CAPM, un modèle d’équilibre dont la validité est démontrée dans un monde espérance écart-type. En retenant un autre modèle d’équilibre, comme celui de Fama et French (1993) ou de Carhart (1997), le succès des stratégies de “*stock picking*” devient tout à fait possible.

Après une étude détaillée des anomalies de marché, il est montré que ces faits empiriques sont plutôt liés aux échecs du CAPM plutôt qu’à des faiblesses de l’efficience, car aucune de ces anomalies n’implique, de notre point de vue, la possibilité de “*market timing*”.

Si ces conclusions vont en faveur de l’efficience informationnelle, nos analyses théoriques ne visent pas à minimiser les contributions de la finance comportementale ou conventionnaliste. L’introduction de la psychologie ou la sociologie en finance a incontestablement amélioré notre compréhension des phénomènes financiers. Cependant, comme les comportements humains disposent d’une complexité et d’une flexibilité extrêmes, leur analyse, voire leur prévision pour anticiper les fluctuations du marché et mettre en place des stratégies permettant de réaliser des rentabilités excessives, semble peu crédible à ce jour.

2. RÉGULARITÉS EMPIRIQUES SUR LES MARCHÉS BOURSIERS

Résumé

Ce chapitre étudie les propriétés statistiques des séries financières. Autrement dit, nous vérifions l'imprévisibilité des rentabilités excessives sur le plan empirique. Si les hypothèses d'auto-corrélation et de mémoire longue sont écartées avec les données journalières et hebdomadaires, les tests BDS indiquent clairement la présence de dépendances non-linéaires. Ainsi, les régularités identifiées dans les moments d'ordre supérieur, ex. la queue épaisse, bouffées de volatilité, multi-échelle, sont présentées dans ce chapitre pour mettre en évidence leur relation avec la soi-disant possibilité de "*market timing*".

Si les modèles économétriques de type ARMA, ARFIMA, STAR ou MSA, caractérisent avec succès certaines séries boursières, ceux-ci ne permettent en aucun cas d'engendrer des rentabilités excessives, car aucun modèle à notre connaissance ne peut prédire le sens des variations futures avec un taux de réussite supérieur à 60%, seuil minimal à atteindre pour qu'une stratégie de "*market timing*" soit rentable (Grabbe 1996).

Introduction

Quelles sont les propriétés empiriques des séries génériques de rentabilités ? Peut-on prévoir la course avec des modèles économétriques ? Ces deux questions intriguent des chercheurs de nombreuses disciplines depuis plus d'un siècle. D'un côté, comprendre la distribution des rentabilités boursières permet de mieux contrôler les risques subis par les investisseurs. D'un autre côté, une prévisibilité des rentabilités futures, au moins dans une certaine mesure, pourrait permettre le "*market timing*", ce qui falsifierait la théorie de l'efficience, ainsi que tous les modèles financiers fondés sur le principe de non-arbitrage.

Ces deux questions sont étudiées dans ce chapitre. En ligne de mire, on montrera la difficulté de battre le "*buy and hold*" malgré la présence de certaines régularités repérables dans les séries financières.

Deux disciplines ont particulièrement contribué à la caractérisation statistique des rentabilités successives.

- Les économètres classiques développent des tests statistiques pour vérifier la présence des régularités. D'une manière générale, ils modélisent les séries financières avec des processus **discrets**. Selon eux, les rentabilités de différentes fréquences peuvent être prises en compte par des modèles distincts.
- Les "écono-physiciens" se concentrent davantage sur la volatilité des rentabilités. Les processus temporels utilisés dans cette discipline sont souvent de nature **continue**, ce qui signifie que les données de toutes les fréquences doivent être caractérisées par un seul modèle.

Dans ce chapitre, nous allons présenter les principales régularités identifiées par ces deux disciplines, notamment les principaux modèles disponibles pouvant les prendre en considération. Un point particulier sera fait sur ce que certains appellent les "faits stylisés" qui caractérisent les chroniques financières.

Selon leur influence sur la prévisibilité des rentabilités futures, les régularités des dynamiques financières peuvent être classées en deux catégories :

1. Les régularités concernant la moyenne des rentabilités (ex. : l'autocorrélation, le phénomène de mémoire longue, la dynamique chaotique). Ces régularités, principalement explorées par les économètres, pourraient rendre possibles les pratiques du "*market timing*".
2. Les régularités portant sur les moments d'ordre supérieur à deux. Ces régularités sont principalement étudiées par les écono-physiciens qui ont considérablement contribué au management des risques financiers.

Or, la validation d'un modèle statistique n'implique pas la possibilité de "*market timing*". Pour démontrer cette assertion, de nombreux auteurs comparent les mo-

dèles ARMA, STAR, MSA ou de neuro-apprentissage¹ avec la stratégie de “*buy and hold*”. En étudiant les principaux résultats obtenus dans ce domaine, nous montrons l'impossibilité de “*market timing*” en nous appuyant sur ces modèles économétriques.

Ce chapitre est organisé en 2 parties :

1. Les propriétés statistiques des rentabilités sont illustrées en conduisant de différents tests empiriques sur des données réelles issues de différents marchés et à différentes fréquences. On distingue dans cette partie, les régularités sur la moyenne de celles sur la volatilité.
2. En s'appuyant sur les régularités identifiées dans la première partie, des modèles économétriques caractérisant les séries financières sont présentés dans la deuxième partie. On s'intéresse alors à leur possible exploitation en terme de “*market timing*”.

2.1 Propriétés statistiques des séries de rentabilités

En considérant l'équation (2.1) comme une représentation très générale des séries financières, caractériser les rentabilités successives revient à déterminer les propriétés statistiques des innovations ε_{t+1} .

$$P_{t+1} = P_t + \varepsilon_{t+1} \quad (2.1)$$

où P_t désigne le prix observé à l'instant t , ε_{t+1} la différence de prix entre les instants t et $t + 1$.

En examinant les cotations successives de 10 marchandises aux Etats-Unis durant une période de 10 ans, King (1930) modélise les ε_{t+1} avec des processus *i.i.d.*. L'équation (2.1) dans ce cas de figure est ce qu'on appelle une “marche au hasard”.

Osborne (1959) propose de remplacer les prix, P_t , de l'équation (2.1) par leur logarithme, $\log(P_t)$. Selon cet auteur, ce sont les différences de logarithmes $\log(P_{n+1}) - \log(P_n)$ qui suivent une loi normale centrée mais pas les différences de prix lui-même. Le modèle “marche au hasard” se transforme ainsi en équation (2.2) :

$$\log(P_{t+1}) = \log(P_t) + \varepsilon_{t+1} \quad (2.2)$$

Où ε_{t+1} est *i.i.d* selon une loi normale centrée².

1. Ces modèles seront présentés dans la section 2.2 page 87.

2. Remarquons que la version continue de l'équation (2.2) est un **mouvement Brownien**,

L'indépendance entre les variations de prix est considérée comme un support de l'efficience informationnelle, car celle-ci implique l'imprévisibilité des rentabilités futures, donc l'impossibilité de battre le "*buy and hold*". Or, en comparant les implications théoriques de l'efficience informationnelle et le modèle "marche au hasard", deux décalages sont à mettre en évidence :

1. **L'absence des rentabilités d'équilibre.** La rentabilité d'un titre donné est composée de deux parties différentes : la rentabilité d'équilibre et la rentabilité excessive. Ce qui est imprévisible selon la théorie de l'efficience est la rentabilité excessive. L'indépendance des rentabilités confirme l'impossibilité de "market timing" donc l'absence des rentabilités excessives. En revanche, la dépendance des rentabilités n'implique pas toujours le rejet de l'efficience, car celle-ci est probablement due à l'évolution des rentabilités d'équilibre.
2. **La différence entre l'imprévisibilité et l'indépendance.** Comme nous l'avons défini dans le premier chapitre, les rentabilités excessives sont dites "imprévisibles" si leur espérance conditionnelle est égale à zéro. Les dépendances observées dans les moments d'ordre supérieur à 2 sont souvent compatibles avec une telle imprévisibilité.

Autrement dit, l'efficience informationnelle demande seulement que les rentabilités excessives soient uncorrélées. Le modèle de "marche au hasard" est une condition suffisante mais non nécessaire à l'absence de rentabilités excessives³.

Motivée par ces deux points, la caractérisation des séries de rentabilités se déroulera en deux étapes : les régularités de la moyenne et celles des moments d'ordre supérieur à deux.

2.1.1 Caractérisation de la moyenne

Trois types de tests sont souvent conduits sur la moyenne des rentabilités : les tests d'autocorrélation, les tests de mémoire longue et les tests de chaos. Dans cette section, ces trois dépendances sont illustrées avec des données intrajournalières, journalières et hebdomadaires : nos résultats montrent l'imprévisibilité des rentabilités futures.

décrit pour la première fois en 1827 par le botaniste Robert Brown en observant des particules à l'intérieur de grains de pollen, puis étudiée par Bachelier (1900) dans sa thèse doctorale en mathématiques, mais officiellement proposée 5 ans plus tard par Einstein, dans ses études sur le mouvement des gaz parfaits.

3. La différence entre une marche au hasard et un jeu équitable est rappelée dans l'annexe J.

I. Tests de dépendance linéaire

La marche au hasard de King (1930) est confirmée par de nombreux travaux empiriques. En étudiant les cours hebdomadaires du blé pendant la période de 1883 à 1934⁴, Working (1934) note que la durée des baisses et des hausses est distribuées selon une loi normale. Ce résultat est généralisé par Roberts (1959) et Fama (1965) aux variations des marchés boursiers. En examinant les rentabilités hebdomadaires du coton, du blé et de 8 autres titres industriels cotés à Londres durant la période 1883-1934, Kendall et Hill (1953) montrent l'absence d'autocorrélation dans ces séries temporelles. Cette même conclusion est obtenue par Alexander (1961) en examinant les différences de prix observées toutes les 2, 4, 8 et 16 semaines sur les marchés Londoniens. Les analyses spectrales confirment aussi l'indépendance des rentabilités successives (Granger et Morgenstern 1963).

Bien que certains auteurs (Brealey 1970, Cunningham 1973, Lo et MacKinlay 1988, Belaire-Franch et Opong 2005) remarquent une autocorrélation dans les rentabilités successives, aucun d'eux ne soutient la possibilité de “*market timing*”, compte tenue de la faiblesse de la corrélation constatée.

Pour vérifier l'absence d'autocorrélation dans les séries de rentabilités, le test “Box Pierce” corrigé de l'hétéroscédasticité⁵, est conduit sur les rentabilités successives des 30 composantes de Dow-Jones. Ce test est mieux adapté aux séries financières que les tests Box-Pierce (Box et Pierce 1970) et Box-Ljung (Ljung et Box 1978), car ces deux derniers ne sont pas robustes aux “sauts” de volatilité souvent observés dans les rentabilités successives (Mignon 1998, Lobato, Nankervis, et Savin 2001).

- Les données utilisées à cet effet sont observées à trois fréquences différentes :
- les rentabilités *tick-by-tick* calculées pour la journée du “05/02/2001”, et extraites de TAQ II⁶.
 - les innovations journalières couvrant la période allant du “01/01/1990” au “22/01/2010” et téléchargées du site Yahoo Finance,
 - et les données hebdomadaires sur la même période d'observation de la même ressource que les rentabilités journalières.

En consultant le tableau de résultat (cf. tableau 2.1 page 60), on remarque trois points :

- Les séries *tick-by-tick* sont négativement autocorrélées sur le premier retard ($lag = 1$). L'hypothèse d'indépendance est rejetée avec un niveau de signification élevé pour la plupart des titres analysés.

4. Sauf 1915-1920.

5. Présenté en détail dans Lobato, Nankervis, et Savin (2001).

6. Une base de données commercialisée par le groupe Euronext-NYSE.

2.1. Propriétés statistiques des séries de rentabilités

Tab. 2.1: Tests de Pierce Box corrigé effectués avec les rentabilités *tick-by-tick* (T), journalières (J) et hebdomadaires (H)

Titre	T (lag=1)	T (lag=10)	J (lag=1)	J (lag=10)	H (lag=1)	H (lag=10)	
1	AA	$3.273e - 06^{***} (-)$	0.2095	0.4711	0.9997	0.7688	1
2	AXP	$3.672e - 09^{***} (-)$	0.1173	0.05875	0.9996	0.5311	0.9997
3	BA	$1.866e - 10^{***} (-)$	0.1108	0.9313	0.9998	0.4005	1
4	BAC	$1.602e - 03^{**} (-)$	0.8058	0.5608	1	0.6652	1
5	CAT	$8.647e - 06^{***} (-)$	0.7522	0.5699	1	0.5083	1
6	CSCO	$0^{***} (-)$	$3.057e - 04^{***} (-)$	$4.91 - 03^{**} (-)$	0.9957	0.2927	0.9998
7	CVX	NA	NA	$9.172e - 03^{**} (-)$	0.9943	$9.95e - 03^{**} (-)$	0.9976
8	DD	$1.0611e - 08^{***} (-)$	0.2400	0.2745	0.9996	$0.02310^* (-)$	0.9998
9	DIS	$4.192e - 05^{***} (-)$	0.2355	0.3367	0.9998	0.4426	0.9999
10	GE	0.578	0.9997	0.6378	0.9998	0.2235	0.999
11	HD	$0^{***} (-)$	$1.181e - 03^{**} (-)$	0.2498	0.9969	0.05815	1
12	HPQ	NA	NA	0.2011	0.9992	0.09312	0.9992
13	IBM	$6.040e - 14^{***} (-)$	$7.256e - 03^{**} (-)$	0.3042	0.9999	0.2518	0.9995
14	INTC	$4.330e - 15^{***} (-)$	$1.876e - 05^{***} (-)$	0.4832	0.9998	0.2435	1
15	JNJ	$2.442e - 09^{***} (-)$	0.13735	0.5422	0.9514	0.1111	0.9966
16	JPM	$1.083e - 08^{***} (-)$	0.3419	0.3156	1	0.1917	0.9998
17	KFT	NA	NA	0.1276	1	0.1505	0.9999
18	KO	0.065	0.9079	0.789	0.9996	0.2402	0.9993
19	MCD	$1.793e - 08^{***} (-)$	0.0823	0.8058	0.9983	0.7948	0.999
20	MMM	$5.083e - 03^{**} (-)$	0.9512	0.1043	0.9946	$0.02684^* (-)$	0.9964
21	MRK	$1.063e - 03^{**} (-)$	0.9823	0.8339	0.9996	0.4158	0.9998
22	MSFT	$2.666e - 10^{***} (-)$	$0.0264^* (-)$	0.1856	0.9999	0.832	0.9997
23	PFE	$3.405e - 06^{***} (-)$	0.1813	0.5189	0.9829	$8.623e - 03^{**} (-)$	0.9958
24	PG	$1.738e - 06^{***} (-)$	0.1718	0.2812	0.9992	$3.565e - 03^{**} (-)$	0.9745
25	T	0	$1.081e - 05^{***} (-)$	0.3409	0.9991	$5.59e - 03^{**} (-)$	0.998
26	TRV	NA	NA	0.1308	1	0.4339	1
27	UTX	$4.773e - 03^{**} (-)$	0.8040	0.1966	0.997	$0.03425^* (-)$	0.9975
28	VZ	$2.845e - 06^{***} (-)$	0.2321	$0.02745^* (-)$	0.9922	$8.709e - 03^{**} (-)$	0.9863
29	WMT	$3.007e - 09^{***} (-)$	0.1266	0.5264	0.9756	$2.048e - 04^{***} (-)$	0.9915
30	XOM	$9.279e - 09^{***} (-)$	$0.0437^* (-)$	$1.517e - 04^{***} (-)$	0.9691	$1.157e - 05^{***} (-)$	0.9273

Hypothèse nulle : absence d'autocorrélation dans la série analysée.
 “*” correspond au seuil de signification à 5%, “**” à 1%, et “***” à 0,1%.
 “(-)” désigne la présence d’une autocorrélation négative.

Ce phénomène est souvent expliqué par le modèle de Roll (1984) fondé sur le rebond de la fourchette de cotation. Ce modèle est étudié dans la suite de cette section pour montrer la compatibilité entre l'impossibilité de "market timing" et l'autocorrélation des rentabilités *tick-by-tick*.

- Parmi les 30 séries hebdomadaires, 10 présentent une autocorrélation significative au seuil de 5%. Ce résultat semble incohérent avec les études antérieures (Fama 1970, Mignon 1998) qui rejettent la présence d'autocorrélation dans les variations hebdomadaires des indices boursiers. Cela peut tenir à l'usage des titres individuels, en lieu et place des indices de marché. Remarquons également que l'autocorrélation constatée reste valable pour une minorité des titres étudiés et sur un intervalle de temps limité. Les profits qu'on pourrait réaliser en exploitant ce phénomène risquent d'être compromis par ce manque de généralité.
- Les autocorrélations observées ne persistent pas dans le temps. En effet, aucune série de rentabilités journalières ou hebdomadaires ne vérifie d'autocorrélation avec un retard élevé ($lag = 10$). Même pour les séries *tick-by-tick*, l'autocorrélation diminue rapidement avec l'augmentation du retard intégré

Parmi ces trois fréquences observés, l'autocorrélation des rentabilités *tick-by-tick* semble mieux établie et plus facile à exploiter par les investisseurs. Cependant, ce phénomène est expliquée par la fourchette de cotation ("*bid-ask spread*") qui empêche le succès des stratégies de type "market timing" (Roll 1984).

En effet, sur les marchés boursiers, il existe toujours un écart entre le prix exigé par le meilleur vendeur ("*ask*") et l'offre proposée par le meilleur acheteur ("*bid*"). Cet écart nommé "fourchette de cotation" évolue en permanence pendant les périodes de transactions.

Selon Roll (1984), les transactions initiées par les vendeurs sont enregistrées aux "*bid*", alors que celles initiées par les acheteurs le sont aux "*ask*".

Soit p_{n-1} , le prix observé d'un actif financier à l'instant $n - 1$, s , la taille de la fourchette de cotation durant la période $n - 1 \rightarrow n + 1$. Sans modifier la fourchette de cotation, les prix peuvent vérifier une différence de s , selon la façon dont les transactions sont initiées.

Les trois prix successifs du titre, notés p_{n-1} , p_n , p_{n+1} , peuvent se trouver dans 8 cas de figure :

2.1. Propriétés statistiques des séries de rentabilités

1. Si p_{n-1} était le “ask” :

p_n et p_{n+1} sont aussi le “ask” : $p_{n-1} \rightarrow p_{n-1} \rightarrow p_{n-1}$

p_n est le “bid” et p_{n+1} le “ask” : $p_{n-1} \rightarrow p_{n-1} - s \rightarrow p_{n-1}$

p_n et p_{n+1} sont tous deux le “bid” : $p_{n-1} \rightarrow p_{n-1} - s \rightarrow p_{n-1} - s$

p_n est le “ask” et p_{n+1} le “bid” : $p_{n-1} \rightarrow p_{n-1} \rightarrow p_{n-1} - s$

2. Si p_{n-1} était le “bid” :

p_n et p_{n+1} sont tous deux le “bid” : $p_{n-1} \rightarrow p_{n-1} \rightarrow p_{n-1}$

p_n est le “ask” et p_{n+1} le “bid” : $p_{n-1} \rightarrow p_{n-1} + s \rightarrow p_{n-1}$

p_n et p_{n+1} sont tous deux le “ask” : $p_{n-1} \rightarrow p_{n-1} + s \rightarrow p_{n-1} + s$

p_n est le “bid” et p_{n+1} le “ask” : $p_{n-1} \rightarrow p_{n-1} \rightarrow p_{n-1} + s$

Sur un marché aléatoire, un titre doit avoir autant de chance d’être échangé au “bid” qu’au “ask”. Chacune des combinaisons précédentes doit avoir autant de chance d’être observée. Les variations de prix successifs, $r_{n-1} = p_n - p_{n-1}$ et $r_n = p_{n+1} - p_n$, peuvent donc se trouver dans les 8 situations suivantes avec la même probabilité pour chacune.

$$\begin{array}{cccc} (0,0) & (-s,s) & (-s,0) & (0,-s) \\ (0,0) & (s,-s) & (s,0) & (0,s) \end{array}$$

Ainsi, la corrélation entre r_n et r_{n-1} est calculée à $-\frac{s^2}{4} < 0$.

L’autocorrélation négative entre les rentabilités *tick-by-tick* est donc expliquée par ce modèle de Roll, fondé sur la présence de la fourchette de cotation. Une telle autocorrélation est difficile à exploiter car les échanges boursiers sont souvent initiés de manière aléatoire.

II. Tests de mémoire longue

Bien que l’autocorrélation des rentabilités ne soit pas significative à court terme, l’effet des rentabilités historiques pourrait persister dans le temps. Autrement dit, si l’impact de la rentabilité d’hier est peu significatif sur celle d’aujourd’hui, son influence peut continuer à exister sur les rentabilités observées dans une semaine, dans un mois, même dans des années.

Ce phénomène nommé “mémoire longue” fut d’abord étudié par l’hydrologue anglais Hurst (1951), en examinant les débits historiques du Nil. L’analyse de

Hurst est introduite en finance par Mandelbrot (1971) pour montrer la différence empirique entre un processus *i.i.d* et les séries de rentabilités réelles.

Depuis cette introduction, on constate une abondance des travaux portés sur l'existence de mémoire longue dans les rentabilités financières. Greene et Fielitz (1977) observent une mémoire longue dans les rentabilités journalières de 200 actions cotées sur le NYSE. Lo (1991) remarque que la méthode R/S⁷ utilisée dans Greene et Fielitz (1977) n'est pas robuste à l'autocorrélation à court terme, il propose donc un R/S rectifié pour tester la mémoire courte contre la dépendance à long terme. En utilisant les données journalières et mensuelles de l'indice CRSP (*Center for Research in Security Prices*), Lo (1991) rejette la présence de mémoire longue. Cette conclusion est confirmée par Ding, Granger, et Engle (1993), Lobato et Robinson (1998) qui soutiennent, en revanche, la présence de mémoire longue dans les rentabilités en valeurs absolues.

Dans la littérature en finance, il existe trois indicateurs de mémoire longue :

1. le coefficient de Hurst, noté H ,
2. l'ordre fractal de différenciation, noté d ,
3. et la pente observée entre la fonction d'autocorrélation et le paramètre de retard ("lag") considéré, noté β .

Plus précisément :

1. Coefficient de Hurst Face à une série temporelle X_t ($t = 1, \dots, n$) d'espérance $\overline{X_n}$, le coefficient de Hurst, noté H , est défini par la statistique R/S , dont R est donné par :

$$R = \text{Max}_{1 \leq k \leq n} \sum_{j=1}^k (X_j - \overline{X_n}) - \text{Max}_{1 \leq k \leq n} \sum_{j=1}^k (X_j - \overline{X_n}) \quad (2.3)$$

et S représente l'écart-type de X_t .

$$S = \left[\frac{1}{n} \sum_{j=1}^n (X_j - \overline{X_n})^2 \right]^{\frac{1}{2}} \quad (2.4)$$

Le coefficient H d'une série donnée varie entre $[0, 1]$. Un bruit gaussien dispose d'un $H = 0,5$. $0,5 < H < 1$ indique la persistance d'autocorrélation, et $0 < H < 0,5$ désigne l'anti-persistance.

2. Ordre fractal de différenciation d Une série X_t est dite "intégrée de l'ordre d ", si

$$(1 - L)^d X_t = \varepsilon_t \quad (2.5)$$

7. Présentée dans le paragraphe "Coefficient de Hurst".

Où ε_t est *i.i.d* d'espérance nulle. L est un opérateur de retard satisfaisant $L \cdot X_t = X_{t-1}$. Le bruit gaussien est intégré de l'ordre $d = 0$. X_t a une mémoire longue si $0 < d < 0,5$, il est anti-persistant si $-0,5 < d < 0$.

3. Descente d'autocorrélation β Soit γ_k la valeur absolue de la fonction d'autocorrélation de X_t , où " k " désigne l'intervalle ("lag") considéré. Si X_t dispose d'une mémoire longue, alors, la relation (2.6) est vérifiée entre γ_k et k .

$$\gamma_k \approx c_k k^{-\beta} \quad (2.6)$$

Où c_k est un paramètre donné, β un nombre réel compris entre $(0, 2)$ qui indique la vitesse à laquelle la corrélation γ_k décroît avec k .

$0 < \beta < 1$ signifie la présence de mémoire longue, $1 < \beta < 2$ indique l'anti-persistance, et $\beta = 1$ est observé pour les bruits gaussiens (Clegg 2005, Mills 1999, p.4, p.114).

Comme tous les indicateurs présentés ci-dessus mesurent le phénomène de mémoire longue, il existe les relations suivantes entre H , d et β (Baillie et Bollerslev 1996, Granger et Hyung 2004).

$$H = 1 - \frac{\beta}{2} \quad (2.7)$$

$$d = H - 0,5 \quad (2.8)$$

$$\beta = 1 - 2d \quad (2.9)$$

Ainsi, on peut distinguer 3 catégories de tests de mémoire longue :

1. La méthode R/S (Mandelbrot 1965) ou R/S corrigé (Lo 1991) qui estime H en appliquant directement la définition de Hurst (1951).
2. L'analyse spectrale (Granger 1966, Geweke et Porter-Hudak 1983, Robinson 1994, Baillie 1996) qui estime le paramètre d en étudiant le périodogramme des séries temporelles.
3. Un autre type d'analyse spectrale estime le paramètre β , cette méthode, nommée également "périodogramme cumulé", est souvent utilisée par les écono-physiciens pour illustrer le phénomène de mémoire longue (Beran et Feng 2001).

Pour tester la mémoire longue, H , d et β sont respectivement estimés pour les 30 composantes de Dow-Jones durant la période du 01/01/1990 au 22/01/2010. En ce qui concerne la fréquence d'observation, seules les données journalières et hebdomadaires sont utilisées, car la non-stationnarité des rentabilités intra-day

peut biaiser la conclusion des tests (Lobato et Robinson 1998, Giraitis, Kokoszka, et Leipus 2001, Taylor 2003).

Notre étude procède en 2 étapes :

1. la stationnarité des séries est vérifiée avec 3 tests différents : le test de Dickey-Fuller augmenté (désormais “DF”) (Dickey et Fuller 1979), le test de Phillips-Perron (désormais “PP”) (Phillips et Perron 1988), et le test de KPSS (Kwiatkowski, Phillips, Schmidt, et Shin 1992). Les résultats des tests sont exposés dans le tableau 2.2 page 66, duquel une conclusion commune émerge : toutes les séries de rentabilités testées sont stationnaires.
2. Après vérification de la stationnarité, les trois indicateurs de mémoire longue sont estimés en appliquant les méthodes suivantes : H est estimé avec l’analyse R/S corrigée de Lo (1991), d avec la méthode de Geweke et Porter-Hudak (1983) (désormais GPH), et β avec la méthode “périodogramme cumulé” de Beran et Feng (2001).

La signification de l’analyse R/S corrigée est vérifiée par le test de Lo (1991). La statistique “ V ” proposée par cet auteur est calculée et exposée pour chaque série dans les colonnes V du tableau 2.3 page 67. Selon Lo (1991), l’hypothèse de la mémoire longue est retenue si $V > 1,862$. Dans le tableau 2.3, les $V > 1,862$ sont mis en *italique*.

La signification des d estimés a fait objet d’un test de Wald. À un niveau de signification de 5%, les d significativement supérieurs à 0 sont mis en *italique*.

La signification de β est vérifiée avec le test du multiplicateur de Lagrange (“*lagrange multiplier*” désormais LM) utilisé par Lobato et Robinson (1998), les résultats indiquant la présence de mémoire longue sont également en *italique* dans les colonnes “ LM ” du tableau 2.3.

Nos tests de mémoire longue permettent de dresser 2 conclusions :

- L’analyse R/S a tendance à surestimer le coefficient de Hurst. Ce résultat est cohérent avec la conclusion de Wallis et Matalas (1970), qui précisent que R/S surestime le coefficient de Hurst si le vrai $H < 0,70$, et le sous-estime si $H > 0,70$.
- Parmi les 60 séries étudiées (30 journalières, 30 hebdomadaires sur les composantes de DJ), seules les rentabilités journalières de l’action “CVX” et “TVR”, et les rentabilités hebdomadaires de “PFE” disposent d’une mémoire longue selon au moins deux tests parmi les trois. Autrement dit, la présence de mémoire longue n’est pas confirmée par nos tests empiriques. Ce résultat est observé par beaucoup d’autres auteurs qui se sont intéressés à ce sujet (Aydogan et Booth 1988, Ding, Granger, et Engle 1993, Lobato et Savin 1998, Granger et Hyung 2004).

2.1. Propriétés statistiques des séries de rentabilités

Tab. 2.2: Tests de stationnarité effectués sur les données journalières (J) et hebdomadaires (H)

	Titre	J (DF)	J (PP)	J (KPSS)	H (DF)	H (PP)	H (KPSS)
1	AA	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
2	AXP	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
3	BA	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
4	BAC	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
5	CAT	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
6	CSCO	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
7	CVX	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
8	DD	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
9	DIS	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
10	GE	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
11	HD	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
12	HPQ	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
13	IBM	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
14	INTC	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
15	JNJ	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
16	JPM	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
17	KFT	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
18	KO	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
19	MCD	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
20	MMM	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
21	MRK	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
22	MSFT	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
23	PFE	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
24	PG	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
25	T	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
26	TRV	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
27	UTX	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
28	VZ	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
29	WMT	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1
30	XOM	< 0.01**	< 0.01**	> 0.1	< 0.01**	< 0.01**	> 0.1

Hypothèse nulle des tests DF et PP : non-stationnarité de la série étudiée.

Hypothèse nulle du test KPSS : stationnarité de la série étudiée.

2.1. Propriétés statistiques des séries de rentabilités

Tab. 2.3: Estimation de H , d et β avec les données journalières et hebdomadaires

Titre	R/S corrigée	Rentabilités journalières				Rentabilités hebdomadaires				
		V	d	β	LM	V	d	β	LM	
1 AA	0.5708	1.6917	0.0545 (0.5545)	0.9923 (0.5039)	0.0859	0.4996	0.9978	0.1010 (0.6010)	1.039 (0.4804)	0.0380
2 AXP	0.6249	2.5279	-0.0643 (0.4357)	0.9686 (0.5157)	1.5693	0.5084	1.0500	0.0062 (0.5062)	0.9997 (0.5002)	0.7096
3 BA	0.6299	2.6223	-0.0804 (0.4196)	1.0679 (0.4660)	0.0013	0.4896	0.9412	-0.1309 (0.3691)	1.1843 (0.4078)	.0084
4 BAC	0.6516	3.081	-0.0252 (0.44748)	1.0062 (0.4969)	0.0980	0.6396	2.2617	-0.0953 (0.4047)	1.0924 (0.4538)	0.9197
5 CAT	0.6396	2.2616	-0.1017 (0.3983)	1.0613 (0.4694)	0.9197	0.5624	1.4406	-0.0907 (0.4093)	1.1901 (0.4049)	1.1089
6 CSCO	0.5671	1.6451	0.0749 (0.5749)	0.8098 (0.5950)	5.2703	0.5854	1.6467	-0.0378 (0.4622)	0.6562 (0.6718)	9.5385
7 CVX	0.6570	3.2065	-0.2000 (0.3000)	1.2672 (0.3664)	6.2738	0.4855	0.9185	-0.1031 (0.3969)	1.2075 (0.3634)	2.3473
8 DD	0.7276	5.416	-0.0332 (0.4668)	1.0793 (0.4603)	0.4220	0.5455	1.3047	0.0722 (0.5722)	1.0686 (0.4657)	1.0606
9 DIS	0.6494	3.0315	-0.0429 (0.4571)	1.0636 (0.4681)	0.0122	0.5201	1.1249	-0.1518 (0.3482)	1.2187 (0.3907)	0.1758
10 GE	0.6896	4.0856	0.0462 (0.5462)	0.9180 (0.5406)	1.3602	0.5908	1.7006	0.2111 (0.7111)	0.6975 (0.6512)	0.1232
11 HD	0.6276	2.5786	-0.0313 (0.4697)	0.9539 (0.5231)	0.0542	0.6243	2.0687	0.1233 (0.6233)	0.7751 (0.6125)	0.0036
12 HPQ	0.6097	2.2567	0.0555 (0.5555)	1.0088 (0.4956)	0.0705	0.5683	1.4904	0.0525 (0.5525)	1.0169 (0.4915)	0.1680
13 IBM	0.7042	4.5533	0.0705 (0.5705)	1.0043 (0.4979)	0.0063	0.6554	2.4811	-0.0410 (0.4590)	0.9369 (0.5316)	1.08427e-06
14 INTC	0.7070	4.6475	0.0462 (0.5462)	0.9142 (0.5429)	1.4581	0.4984	0.9908	0.0774 (0.5774)	0.9270 (0.5365)	0.3591
15 JNJ	0.6408	2.8432	0.0227 (0.5227)	1.0619 (0.4690)	0.9137	0.6121	1.9258	-0.0155 (0.4845)	0.9935 (0.5033)	1.2313
16 JPM	0.6554	3.1681	0.0199 (0.5200)	1.0744 (0.4628)	0.1791	0.6015	1.8104	-0.1155 (0.3845)	1.0493 (0.4753)	1.4451
17 KFT	0.4861	0.9128	-0.2478 (0.2522)	1.3229 (0.3385)	0.2955	0.6437	2.0545	-0.3805 (0.2255)	1.3367 (1.3367)	1.5040
18 KO	0.5799	1.8100	0.0121 (0.5121)	1.0180 (0.4910)	1.1227	0.5039	1.0232	0.3374 (0.8374)	0.9506 (0.5247)	0.0065
19 MCD	0.7186	5.0661	0.0904 (0.5904)	1.0214 (0.4893)	0.7579	0.5086	1.0516	-0.0187 (0.4813)	1.0230 (0.4885)	0.0008
20 MMM	0.7114	4.8010	-0.0787 (0.4213)	1.1864 (0.4068)	2.2533	0.4561	0.7736	-0.0159 (0.4841)	1.2305 (0.3847)	3.1075
21 MRK	0.5105	1.0814	0.0496 (0.5496)	0.9649 (0.5175)	0.0028	0.7119	3.4528	0.1040 (0.6040)	0.8250 (0.5875)	0.6468
22 MSFT	0.6432	2.8953	0.0903 (0.5903)	0.9706 (0.5147)	0.1150	0.5545	1.3756	0.1235 (0.6235)	0.7275 (0.6362)	0.8557
23 PFE	0.6504	3.0541	0.0051 (0.5051)	0.9240 (0.5380)	0.0178	0.6630	2.594	0.2667 (0.7667)	.6506 (0.6747)	0.5745
24 PG	0.6501	<i>3.0463</i>	-0.0350 (0.4650)	1.0993 (0.4503)	3.4726	0.5580	1.4040	0.1304 (0.6304)	1.0907 (0.4546)	1.3810
25 T	0.5045	1.0336	-0.0960 (0.4040)	1.0790 (0.4605)	2.4319	0.4587	0.7853	0.0703 (0.5703)	0.8521 (0.5740)	0.3060
26 TRV	0.6213	2.4608	-0.1619 (0.3381)	1.2428 (0.3786)	5.0381	0.5300	1.1919	-0.0587 (0.4413)	1.2037 (0.3981)	0.3913
27 UTX	0.6651	3.4049	-0.0422 (0.4578)	1.1129 (0.4435)	0.8331	0.4651	0.8154	-0.0618 (0.4382)	1.0780 (0.4610)	1.5594
28 VZ	0.6101	2.2643	-0.1382 (0.3618)	1.1388 (0.4306)	0.7172	0.4649	0.8146	-0.0773 (0.4227)	1.0960 (0.4520)	2.7287
29 WMT	0.7976	9.1099	-0.0310 (0.4690)	1.0261 (0.4869)	0.1196	0.5797	1.5939	0.0692 (0.5692)	0.8203 (0.5898)	0.3781
30 XOM	0.5148	1.1158	-0.0523 (0.4477)	1.1671 (0.4165)	3.3948	0.4911	0.9493	0.0763 (0.5763)	0.9785 (0.5107)	0.8167

- Interprétation de la valeur V : selon les statistiques fournies par Lo (1991), avec un niveau de signification à 5%, l'hypothèse nulle $H = 0,5$ est acceptée si $V \in [0.809, 1.862]$. Les valeurs de V indiquant $H > 0,5$ sont mises en *italique*.
- La signification des d estimés est vérifiée avec le test de Wald, les d significativement supérieurs à 0 au seuil de 5%, sont mis en *italique*.
- La signification de β est vérifiée avec le test LM de Lobato et Robinson (1998). La largeur de bande utilisée à cet effet est " $m = n^{0.6}$ ". Les résultats du LM indiquant $\beta < 1$ sont mis en *italique*.

III. Vers une dépendance non linéaire

Bien que les tests d'autocorrélation et de mémoire longue ne décèlent aucune preuve solide de dépendance financière, certaines structures non linéaires peuvent toujours exister dans les innovations de prix successives.

Dans cette section, le test de Brock, Scheinkman, Dechert, et LeBaron (1996) (désormais BDS, dont les détails sont présentés dans l'annexe F), sensible aussi bien aux dépendances linéaires qu'aux structures non linéaires, est utilisé pour vérifier l'absence des structures non linéaires.

Ici, on teste toujours les séries hebdomadaires et journalières présentées dans la section 2.1.1 page 59. Les rentabilités *tick-by-tick* ne sont pas soumises au test BDS, car leur autocorrélation négative rendrait difficile l'interprétation des résultats obtenus. De plus, comme le test BDS est sensible à la non-stationnarité des séries analysées, les données *tick-by-tick* qui ont "échoué" au test de stationnarité sont inadaptées.

En consultant le tableau 2.4 page 69, on remarque que l'hypothèse de distribution *i.i.d.* est rejetée pour toutes les séries étudiées. Même si les rentabilités journalières de l'action "KO" et les variations hebdomadaires de "KFT", "MRK" et "PFE" délivrent des "p-value" légèrement supérieures à 5%, en augmentant la dimension m et la distance critique ε , on rejette rapidement l'hypothèse nulle.

Comme le test BDS vérifie l'hypothèse *i.i.d.* contre :

1. les dépendances stochastiques linéaires (l'hypothèse écartée par les tests de corrélation),
2. les dépendances stochastiques non linéaires (ex. : les modèles GARCH, STAR),
3. les dépendances déterministes non linéaires (i.e. le chaos),

deux types de structures non linéaires sont susceptibles d'exister à l'issue de cette analyse : la dépendance stochastique non linéaire et le chaos.

Un domaine de recherche consiste donc à distinguer ces deux types de dépendances. La section prochaine a pour objectif de récapituler les principaux résultats obtenus à ce sujet en soulignant leur impact sur la prévisibilité des rentabilités futures.

2.1. Propriétés statistiques des séries de rentabilités

Tab. 2.4: Tests BDS effectués avec les rentabilités journalières et hebdomadaires

Titre	Rentabilités journalières				Rentabilités hebdomadaires			
	$\epsilon = 0.0124$	$\epsilon = 0.0247$	$\epsilon = 0.0371$	$\epsilon = 0.0495$	$\epsilon = 0.0124$	$\epsilon = 0.0247$	$\epsilon = 0.0371$	$\epsilon = 0.0495$
1 AA	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
2 AXP	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
3 BA	(0,0)	(0,0)	(0,0)	(0,0)	<i>(0.0250,0.0018)</i>	<i>(0.0091,0.0002)</i>	<i>(2e-04,0e+00)</i>	(0,0)
4 BAC	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
5 CAT	(0,0)	(0,0)	(0,0)	(0,0)	<i>(0e+00,1e-04)</i>	<i>(1e-04,0e+00)</i>	<i>(1e-04,0e+00)</i>	(0,0)
6 CSCO	(0,0)	(0,0)	(0,0)	(0,0)	<i>(4e-04,0e+00)</i>	(0,0)	(0,0)	(0,0)
7 CVX	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
8 DD	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
9 DIS	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
10 GE	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
11 HD	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
12 HPQ	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
13 IBM	(0,0)	(0,0)	(0,0)	(0,0)	<i>(3e-04,0e+00)</i>	<i>(0.0018,0.0000)</i>	<i>(1e-02,1e-04)</i>	<i>(0.0129,0.0011)</i>
14 INTC	(0,0)	(0,0)	(0,0)	(0,0)	<i>(0.0018,0.0001)</i>	<i>(0.0019,0.0000)</i>	<i>(0.0051,0.0001)</i>	<i>(0.0074,0.0002)</i>
15 JNJ	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	<i>(2e-04,0e+00)</i>	<i>(1e-04,0e+00)</i>	(0,0)
16 JPM	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
17 KFT	(0,0)	(0,0)	(0,0)	(0,0)	<i>(0.128,0.047)</i>	<i>(0.0690,0.0069)</i>	<i>(0.0246,0.0019)</i>	<i>(0.0376,0.0033)</i>
18 KO	<i>(0.128,0.047)</i>	<i>(0.0690,0.0069)</i>	<i>(0.0246,0.0019)</i>	<i>(0.0376,0.0033)</i>	(0,0)	(0,0)	(0,0)	(0,0)
19 MCD	(0,0)	(0,0)	(0,0)	(0,0)	<i>(0.0435,0.0284)</i>	<i>(0.0073,0.0030)</i>	<i>(0.0097,0.0021)</i>	<i>(0.0217,0.0035)</i>
20 MMM	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
21 MRK	(0,0)	(0,0)	(0,0)	(0,0)	<i>(0.0882,0.0160)</i>	<i>(0.0648,0.0102)</i>	<i>(0.0433,0.0071)</i>	<i>(0.0340,0.0065)</i>
22 MSFT	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	<i>(1e-04,0e+00)</i>
23 PFE	(0,0)	(0,0)	(0,0)	(0,0)	<i>(0.1771,0.0027)</i>	<i>(0.0885,0.0029)</i>	<i>(0.0449,0.0010)</i>	<i>(0.0351,0.0012)</i>
24 PG	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
25 T	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
26 TRV	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
27 UTX	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
28 VZ	(0,0)	(0,0)	(0,0)	(0,0)	<i>(1e-04,0e+00)</i>	(0,0)	(0,0)	(0,0)
29 WMT	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
30 XOM	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)	(0,0)

Hypothèse nulle : la série est *i.i.d.* Ce tableau expose les “p-values” des tests BDS conduits sur les rentabilités journalières et hebdomadaires. Chaque case du tableau est composée par deux valeurs entre parenthèses “(a,b)” où “a” représente le “p-value” obtenu en retenant une dimension de plongement (“*embedding dimension*”) $m = 2$, et “b” celui obtenu avec $m = 3$. Ici, les résultats significatifs à 5% sont mis en *italique*. On remarque que l’hypothèse nulle est rejetée pour pratiquement toutes les séries de rentabilités examinées.

Le chaos et l'imprévisibilité des rentabilités futures

Un système déterministe ne produit pas toujours une dynamique simple. Tel est le principe du phénomène de chaos. Ce phénomène, initialement découvert par le célèbre mathématicien français Henri Poincaré à la fin du 19ième siècle, a suscité un grand intérêt dans les années 60.

L'application de la théorie du chaos en finance date de la fin des années 80, peu de temps après le krach en 1987. En effet, le 19 octobre 1987, l'indice Dow-Jones industriel a enregistré en une seule journée de transactions, un choc négatif de $-22,6\%$. Ce constat sans précédent a amplement compromis la capacité des modèles stochastiques classiques à décrire les dynamiques financières. La nature des séries financières (déterministe ou stochastique) devient ainsi un domaine de recherche captant l'attention des chercheurs en finance.

Basé sur une série de 5200 rentabilités journalières, Scheinkman et LeBaron (1989) n'ont pas constaté de preuve solide soutenant la nature chaotique des séries financières. En testant l'hypothèse du chaos contre de nombreux processus stochastiques, Hsieh (1991) associe le rejet de l'hypothèse de distribution *i.i.d.* à la prévisibilité des variances ("*conditional heteroscedasticity*"). Ce résultat est en ligne avec la conclusion de Diebold et Nason (1990) qui remettent en question, avec un test non paramétrique, la capacité des modèles non linéaires à prédire les taux de change. En étudiant les taux de change entre la livre et 10 autres monnaies étrangères, Brooks (1998) rejette l'hypothèse de dynamique chaotique. Ce rejet est confirmé par Mignon (1998) en examinant les principaux indices des 7 marchés industriels. Plus récemment, Sosvilla-Rivero, Fernández-Rodríguez, et Andrada-Félix (2005) détectent une dynamique chaotique dans les taux "franc/dollar", tout en soulignant que celle-ci n'est observée que dans un seul des 3 échantillons analysés et uniquement pendant la sous-période 1970-1980.

Dans cette section, nous allons présenter brièvement le concept de chaos déterministe, ses définitions et ses propriétés mathématiques, qui a suscité de grands débats de nature philosophique sur la notion de hasard et d'imprévisibilité (Werndl 2009).

La théorie du chaos étudie le comportement des systèmes dynamiques. Un système dynamique est un modèle mathématique qui contient un ensemble d'états possibles X et une équation d'évolution qui décrit comment ces états se transforment dans le temps (cf. équation 2.10).

$$x_{n+1} = T(x_n) \tag{2.10}$$

où $x_n \in X$, $n \in \mathbb{N}$.

Par exemple, l'équation 2.11, nommée équation logistique, est souvent utilisée pour illustrer le comportement des systèmes dynamiques.

$$x_{n+1} = wx_n(1 - x_n) \quad (2.11)$$

où w est un paramètre donné, et w, x_n satisfont :

$$\begin{cases} 0 \leq w \leq 4 \\ x_n \in [0, 1] \end{cases}$$

Le comportement de cette équation varie en fonction de w :

- Quand $0 \leq w \leq 2$, la dynamique du système est illustrée par la figure 2.1(a) page 72. Quelle que soit la valeur initiale x_0 , le système (x_n) converge toujours vers la solution $x^* = \frac{w-1}{w}$.
- Quand $2 < w \leq 3$, la dynamique du système est illustrée par la figure 2.1(b). Bien que le système oscille autour de la solution $x^* = \frac{w-1}{w}$, la distance entre x_n et x^* diminue continuellement. Le système reste donc convergent.
- Quand $3 < w \leq 4$ la dynamique du système est illustrée par la figure 2.1(c) : le système ne converge pas vers une valeur constante, la distance entre x_n et x^* peut varier considérablement d'un moment à l'autre. Bien que les deux valeurs initiales sont relativement proches (0,200 et 0,201), les deux trajectoires engendrées sont remarquablement distinctes. De plus, leur différence ne s'affaiblit pas dans le temps. On constate en effet un comportement de chaos.

Dans cet exemple, on peut constater les deux principales caractéristiques des systèmes chaotiques : la présence d'attracteur et "effet papillon".

- La présence d'attracteur : $x^* = \frac{w-1}{w}$ est l'attracteur du phénomène chaotique constaté. Comme on peut le remarquer dans la figure 2.1(c), les x_n tournent autour de x^* comme si ce dernier exerçait une attirance irrésistible pour tous les points du système.

En termes mathématiques, sans converger vers x^* , le système X satisfait :

$$\liminf_{n \rightarrow \infty} (x_n - x^*) = 0 \quad (2.12)$$

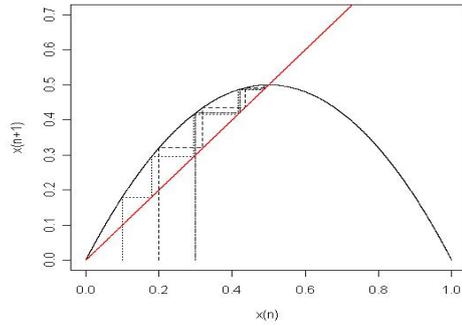
- "L'effet papillon" (ou la sensibilité aux conditions initiales) : Selon Mignon (1998), un système X est dit "sensible" aux conditions initiales, s'il existe un nombre $r > 0$ tel que pour tout point $x, y \in X$ et $\varepsilon > 0$, il existe :

- $d(x, y) < \varepsilon$,
- et un entier $n > 0$, tel que $d(T^n(y), T^n(x)) \geq r$.

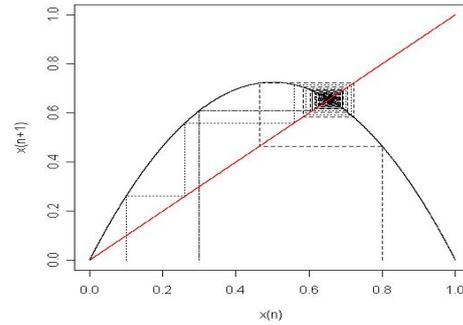
Ici, $d(x, y)$ désigne la distance entre les deux points initiaux x et y . $T^n(x)$ représente la nième itération de x (souvent notée x_n).

Ce phénomène est bien illustré par la figure 2.1(c) : avec une différence en valeurs initiales de $0,201 - 0,200 = 0,001$, les deux trajectoires chaotiques s'écartent considérablement dans le temps.

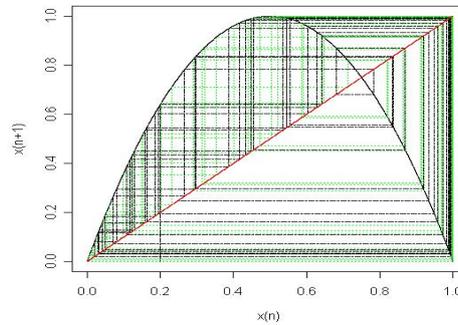
2.1. Propriétés statistiques des séries de rentabilités



(a) Dynamique de l'équation logistique avec $w = 2$. Ici, la courbe quadratique correspond à la relation $x_n \sim x_{n-1}$. La ligne droite désigne $y = x$. Les lignes discontinues montrent les trajectoires de x_n avec les valeurs initiales $x_0 = \frac{1}{10}, \frac{2}{10}$ et $\frac{3}{10}$.



(b) Dynamique de l'équation logistique avec $w = 2,9$. Ici, la courbe quadratique correspond à la relation $x_n \sim x_{n-1}$. La ligne droite désigne $y = x$. Les lignes discontinues montrent les trajectoires de x_n avec les valeurs initiales $x_0 = \frac{1}{10}, \frac{3}{10}$ et $\frac{8}{10}$.



(c) Dynamique de l'équation logistique avec $w = 4$. Ici, La courbe quadratique correspond à la relation $x_n \sim x_{n-1}$. La ligne droite désigne $y = x$. La ligne discontinue en vert désigne la trajectoire de X avec $x_0 = 0,200$, celle en noir est la trajectoire avec $x_0 = 0,201$.

Fig. 2.1: Dynamique de l'équation logistique, différentes valeurs de w .

“L’effet papillon” est particulièrement intéressant, car il implique l’imprévisibilité des processus chaotiques : toute erreur dans la spécification d’un point initial peut conduire à une orbite dont le comportement de long terme n’a aucune ressemblance avec l’orbite initiale. De plus, “en pratique, on ne peut jamais spécifier l’état d’un système naturel avec une précision mathématique parfaite”⁸, il s’ensuit que cette erreur, aussi insignifiante soit-elle, va se trouver amplifiée et altérer le comportement macroscopique du système (Mignon 1998, p.154). “L’effet papillon” interdit donc la prévision à long terme des systèmes chaotiques.

Quant à la prévision à court terme, cela semble possible si la dynamique du système (la fonction $T(\cdot)$ dans l’équation 2.10) est parfaitement connue d’avance. Cette condition, accessible en théorie mais peu réaliste en pratique, réduit sensiblement la prévisibilité à court terme des systèmes chaotiques. Bien que nous puissions tester la présence de chaos dans les séries financières, retrouver l’équation $x_{n+1} = T(x_n)$ à partir des x_n observés reste extrêmement difficile, voire impossible.

Ainsi, si les variations de prix sont réellement dirigées par une dynamique chaotique, les rentabilités futures seront totalement imprévisibles, battre le “*buy and hold*” avec une stratégie “*market-timing*” sera donc impossible.

Contrairement aux sections précédentes, ici, les tests de chaos ne sont pas conduits sur notre base de données pour 3 raisons différentes :

1. La présence de chaos est compatible avec l’impossibilité de “*market timing*”. Donc, ce qui nécessite une discussion, c’est plutôt l’absence de chaos dans les séries financières, car c’est uniquement dans ce cas de figure que les modèles statistiques pourraient prédire les tendances du marché.
2. Comme nous l’avons évoqué en début de section, la plupart des tests réalisés rejettent l’hypothèse du système chaotique.
3. Le chaos est un cas spécifique de la dépendance non linéaire qui sera étudiée en détail dans la section 2.2.3 (intitulée “*modèles multi-états*”). Selon Hsieh (1991),

“If x_n is a chaotic process, it can be written as :

$$x_n = T(x_{n-1}, \dots) \tag{2.13}$$

This is a special case of a more general category of nonlinear process :

$$x_n = T(x_{n-1}, \dots) + \varepsilon_n \tag{2.14}$$

8. En effet, les ordinateurs ne travaillent jamais sur les objets mathématiques continus, ils sont toujours restreints à choisir un degré de prévision raisonnable. Ce point sera développé dans le chapitre suivant lorsque la complexité des séries financières sera estimée avec les algorithmes de compression.

where ε_n satisfies the condition that $E(\varepsilon_n|x_{n-1}) = 0$ " (Hsieh 1991, p.1860)

Pour résumer, à ce stade d'analyse, 3 conclusions émergent de nos études empiriques :

1. les rentabilités intra-journalières sont négativement autocorrélées, ce phénomène, probablement expliqué par la présence de la fourchette de cotation, demeure compatible avec l'impossibilité de "*market timing*".
2. d'une manière générale, les rentabilités journalières et hebdomadaires forment des séries stationnaires et non autocorrélées. Cependant, le test BDS a révélé une dépendance non linéaire dans les séries examinées.
3. Trois types de dépendances non linéaires sont exploitées dans la littérature en finance : l'autocorrélation de la volatilité (souvent prise en compte par les modèles de famille GARCH), la dynamique chaotique qui implique l'imprévisibilité des rentabilités futures et la dépendance stochastique non linéaire (souvent interprétée par les modèles multi-états).

Après avoir étudié les propriétés statistiques de la moyenne des rentabilités, nous allons étudier, dans la prochaine section, les caractéristiques de leurs moments d'ordre supérieur à deux.

2.1.2 Propriétés des moments d'ordre supérieur à deux

Bien que les moments d'ordre supérieur à deux n'aient pas toujours un impact sur la prévisibilité de la moyenne, ceux-ci sont essentiels pour caractériser et modéliser les séries financières. Certains travaux en simulation de marché utilisent leurs propriétés comme l'objectif de calibrations (Lux et Schornstein 2005, Derveeuw, Beaufils, Brandouy, et Mathieu 2007). Beaucoup d'études en management des risques sont également fondées sur la caractérisation de la volatilité (Bouchaud et Potters 2000).

Contrairement aux dépendances de la moyenne, les régularités identifiées dans les moments d'ordre supérieur sont souvent estimées invariantes à la fréquence, la période et le marché d'observation. Ces régularités, nommées "faits stylisés", ont été observées dans de nombreux travaux empiriques, elles sont aujourd'hui admises par la plupart des chercheurs en finance.

Par exemple, Mandelbrot (1963) et Fama (1965) constatent que des valeurs extrêmes sont plus souvent observées dans les rentabilités financières que dans une distribution normale. Il s'agit du phénomène dit de queue épaisse ("*fat tail*"). Les auteurs cités ci-dessus remarquent également que l'autocorrélation des rentabilités absolues est significative, ce phénomène étant lié aux bouffées de volatilité ("*volatility clustering*"). En étudiant les rendements des actions, Granger et Ding (1994)

notent que l'autocorrélation de la série temporelle $|r_1|^m, |r_2|^m, \dots, |r_t|^m, m > 0$ (où r_t est la variation de prix à l'instant t et m est un paramètre constant) atteint son maximum quand $m = 1$. Il s'agit de ce qu'on appelle l'effet de Taylor (“*Taylor effect*”). Convaincus par la généralité de la loi exponentielle, Müller, Dacorogna, Olsen, Pictet, Schwartz, et Morgeneegg (1990) observent dans un optique similaire, un phénomène d'échelle dans les séries financières (“*scaling laws*”). Bouchaud et Potters (2000) constatent une corrélation négative entre la variation passée et la volatilité future, nommée l'effet de levier (“*leverage effect*”).

Les synthèses réalisées par Mills (1999), Cont (2001), et Bouchaud et Potters (2000) proposent des illustrations graphiques de l'ensemble des faits stylisés observés⁹. Ma (2009) vérifie l'invariance de ces régularités par rapport à la fréquence, la période et le marché d'observation.

Étant donné le nombre des travaux consacrés à ce sujet, nous nous concentrons ici sur les faits stylisés qui ont des impacts potentiellement forts sur la modélisation des séries financières (cf. la présence excessive des valeurs extrêmes, les bouffées de volatilité et le phénomène multi-échelle). De plus, au lieu de proposer des illustrations graphiques, nous insistons, dans cette section sur les indicateurs quantitatifs des faits stylisés permettant de mesurer la similarité entre une série financière et un modèle statistique (Veryzhenko, Brandouy, et Mathieu 2010).

I. Phénomène de queue épaisse

Pendant presque deux siècles, la loi normale a été considérée comme “la loi universelle des erreurs (Mandelbrot 1997, 92-96)”. C'est uniquement dans les années 50 que des physiciens (par exemple, Davis (1952), Epstein et Sobel (1953)) remarquent que “l'usage systématique de cette loi est dû à un énorme malentendu entre les mathématiciens qui l'étudient pour satisfaire les besoins du monde physique et les physiciens qui l'utilisent pour profiter des résultats obtenus en mathématiques (Mandelbrot 1997, 92-96)”.

Ce même progrès a été réalisé en finance par Mandelbrot (1963) et Fama (1965), en découvrant le phénomène de “queue épaisse” ou de “kurtosis excessif”.

Littéralement, le phénomène de “queue épaisse” se définit par une fréquence trop élevée des valeurs extrêmes (rentabilités positives ou négatives s'écartant

9. Certains auteurs, notamment Cont (2001), considèrent l'effet de calendrier (section 1.3) comme un fait stylisé. Or, selon notre distinction entre les anomalies du marché et les faits stylisés, l'effet de calendrier est classé avec les anomalies de marché, car leur identification nécessite plus de données que les rentabilités historiques.

exagérément de la moyenne). Ce phénomène illustré dans la figure 2.2 est traditionnellement caractérisé par la statistique de “*kurtosis*” ou quatrième moment normalisé¹⁰ des rentabilités (Fama 1965, Mandelbrot 1963, Wasserfallen et Zimmerman 1985).

La loi normale ayant une *kurtosis* égal à 3, un échantillon possède une queue épaisse si sa *kurtosis* est supérieur à 3.

En revanche, le quatrième moment d’une distribution donnée n’existe pas toujours : une présence excessive des valeurs extrêmes peut entraîner la non-convergence de la *kurtosis*. Utiliser cet indicateur comme la mesure de la queue épaisse pourrait réduire son champ d’application.

À cet effet, l’indice de Pareto semble mieux convenir (Shiryaev 1999).

Soit X , une variable aléatoire donnée, son indice de Pareto est défini par le paramètre “ α ” de l’équation (2.16) :

$$P[X > x] \sim x^{-\alpha} \quad k < x < \infty \quad (2.16)$$

où k désigne la frontière entre les valeurs “normales” et “extrêmes”. L’indice α ainsi défini mesure la vitesse à laquelle la probabilité $P(X > x)$ diminue avec x .

La convergence du s ième moment normalisé de X dépend de son indice de Pareto. Par définition, le s ième moment de X est calculé par

$$E(X^s) = \int_k^\infty P[X = x] \cdot x^s dx \sim \int_k^\infty (-\alpha x^{-\alpha-1} \cdot x^s) dx \quad (2.17)$$

Ainsi,

- pour tout $s > \alpha$, l’équation (2.17) tend vers ∞ , le s ième moment ne converge pas.
- pour tout $s < \alpha$, le s ième moment converge vers une valeur finie.

Par exemple, l’indice Pareto d’une loi normale est ∞ , ce qui implique la convergence de tous ses moments normalisés. Ayant un indice de Pareto entre 0 et 2, la loi Pareto(-Levy) stable¹¹ dispose d’une variance infinie.

Bien entendu, la convergence des moments normalisés reste un point technique en modélisation des séries financières. Elle n’a pas d’effet direct sur l’exploitation

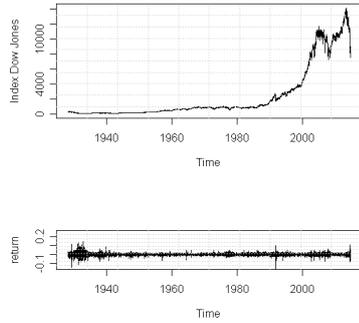
10. La *kurtosis* d’une variable ε est défini comme suit :

$$k(\varepsilon) = E \frac{(\varepsilon - \mu)^4}{\sigma^4} \quad (2.15)$$

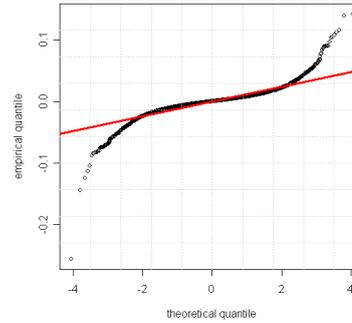
où μ désigne la moyenne et σ l’écart-type de ε .

11. Introduite en finance par Mandelbrot (1963) pour caractériser les rentabilités successives.

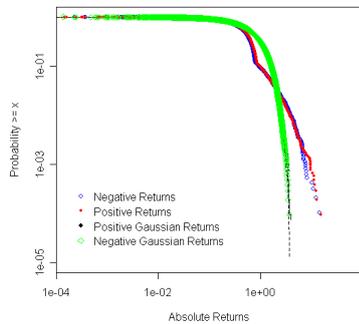
2.1. Propriétés statistiques des séries de rentabilités



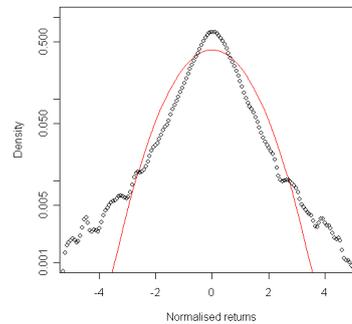
(a) La courbe du haut représente la valeur de l'indice DJ observé à chaque clôture journalière, celle du bas désigne ses variations successives (rentabilités).



(b) Q-Q plot des variations observées. La ligne droite est ce que nous devrions observer si les rentabilités étaient normalement distribuées. Le grand "S" constaté est expliqué par une fréquence trop élevée des valeurs extrêmes.



(c) Fonction de survie. L'axe "Y" représente $\log(P(X > x))$, et "X", $\log(x)$. On constate qu'il est plus probable d'observer des valeurs extrêmes dans les données de DJ que dans les réalisations d'une loi normale.



(d) La courbe continue représente la densité de la loi normale, les points noirs désignent celle des rentabilités observées. Encore une fois, on constate une présence excessive des valeurs extrêmes dans les variations de DJ.

Fig. 2.2: Illustration qualitative du phénomène de queue épaisse. Données utilisées : variations journalières de l'indice DJ industriel durant la période du 10/01/1928 au 01/13/2009.

potentielle des cours historiques. Cependant, l'existence des premier et deuxième moments est essentielle pour le fondement du CAPM ainsi que tous les autres modèles d'équilibre qui examinent l'espérance et la variance des rentabilités.

Pour mesurer “la queue” des rentabilités financières, l'indice de Pareto est estimé pour les données présentées dans la section 2.1.1 page 59. Comme Dacorogna et al. (2001) l'ont réalisé, pour chaque fréquence étudiée, nous réunissons les rentabilités de tous les titres observés sans tenir compte de la notion du temps. La méthode d'estimation est celle de Hill (Hill 1975) dont l'efficacité a été bien étudiée par Hall (1982), Mason et Turova (1994) et Dacorogna, Muller, Pictet, et Vries (1995).

Les résultats obtenus, exposés dans la figure 2.3 page 79, peuvent être résumés en trois points :

- L'indice de Pareto est relativement stable quand il s'agit des données à haute fréquence (*tick-by-tick* et journalière). Celui des rentabilités hebdomadaires s'est avéré bien plus élevé, car leur distribution est plus proche d'une loi normale.
- L'indice de Pareto des rentabilités négatives est inférieur à celui des positives. Cela signifie que les valeurs extrêmes sont plus souvent négatives que positives. Cette asymétrie, notée également par Fama (1965) et Sornette (2002), est expliquée par la présence des crises financières finalement assez fréquentes.
- Les indices de Pareto estimés étant supérieurs à 3, les trois premiers moments normalisés devraient exister pour les rentabilités financières. Cette conclusion est aussi obtenue par Mills (1999, p.188-192) en utilisant les variations journalières des indices FT 30 et SP500, puis confirmée par Jansen et Vries (1991), L. Haan et Vries (1991) et Loretan et Philipps (1994) en vérifiant les données des marchés monétaires. Comme McCauley (2006) l'a résumé :

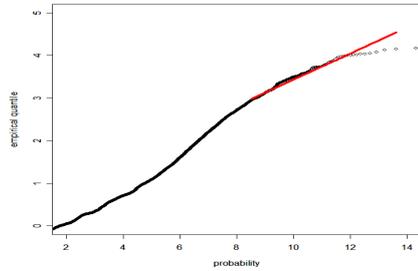
“There is apparently no universality of scaling exponents in finance, where the fat tail exponents range from roughly 2 to 7 and are market dependent.” McCauley (2006, 10)

Après avoir vérifié le phénomène de “queue épaisse”, nous allons étudier, dans la prochaine section, une autre caractéristique des rendements boursiers, la présence de mémoire longue dans la volatilité.

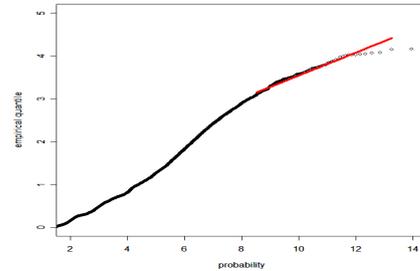
II. Présence de mémoire longue dans la volatilité

Le phénomène des bouffées de volatilité (*volatility clustering*) se traduit par la présence des périodes agitées dans les séries financières suivies de périodes moins

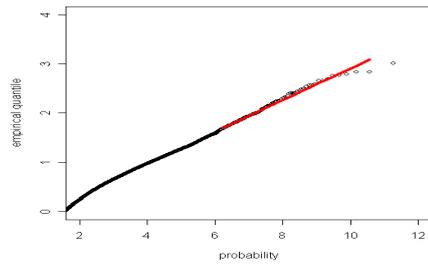
2.1. Propriétés statistiques des séries de rentabilités



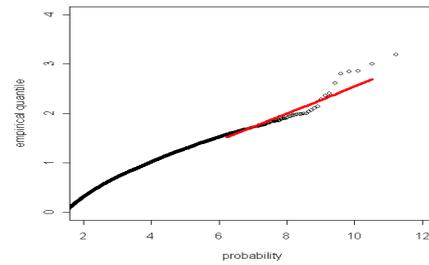
(a) Rentabilités négatives *tick-by-tick* : la pente de la ligne droite étant estimée à 0.305, l'indice de Pareto $\alpha = 1/0.305 = 3.28$



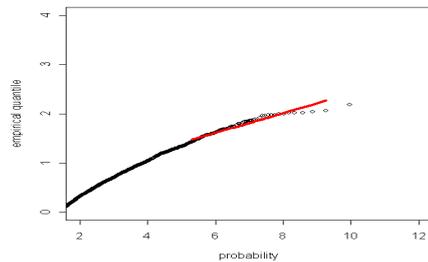
(b) Rentabilités positives *tick-by-tick* : la pente de la ligne droite étant estimée à 0.267, l'indice de Pareto $\alpha = 1/0.267 = 3.74$



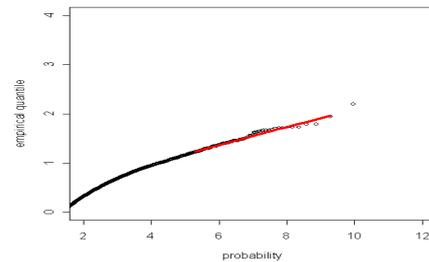
(c) Rentabilités négatives journalières : la pente de la ligne droite étant estimée à 0.321, l'indice de Pareto $\alpha = 1/0.321 = 3.12$



(d) Rentabilités positives journalières : la pente de la ligne droite étant estimée à 0.274, l'indice de Pareto $\alpha = 1/0.274 = 3.65$



(e) Rentabilités négatives hebdomadaires : la pente de la ligne droite étant estimée à 0.2025, l'indice de Pareto $\alpha = 1/0.2025 = 4.94$

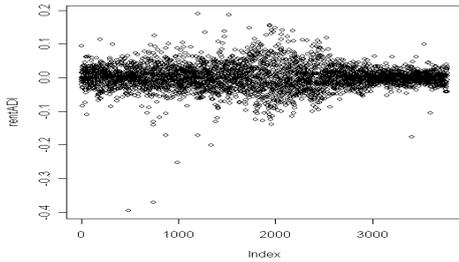


(f) Rentabilités positives hebdomadaires : la pente de la ligne droite étant estimée à 0.1858, l'indice de Pareto $\alpha = 1/0.1858 = 5.38$

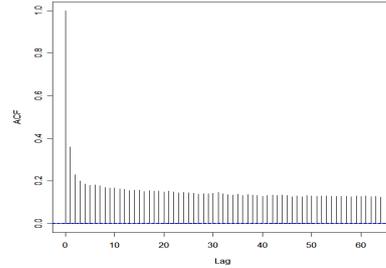
Fig. 2.3: L'estimation de l'indice de Pareto

2.1. Propriétés statistiques des séries de rentabilités

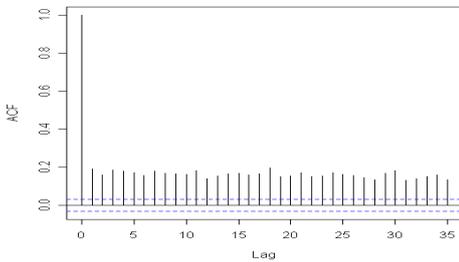
volatiles (cf. figure 2.4(a)). Statistiquement, ce phénomène implique l'autocorrélation des rentabilités absolues. La fonction d'autocorrélation (désormais ACF) des rentabilités absolues est une illustration directe de ce fait stylisé (cf. figure 2.4).



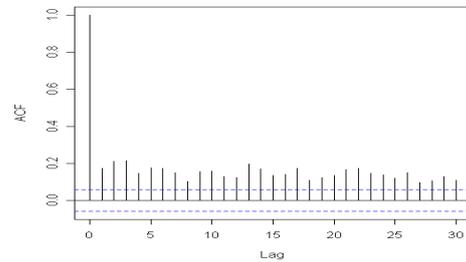
(a) Rentabilités journalières de l'action Analog Devices durant la période du 01/01/1994 au 01/01/2008. On constate ici qu'entre les 1500ième et 2500ième observations, ce titre est plus volatile que pendant le reste du temps, cette zone de perturbation plus importante est ce qu'on appelle "une bouffée de volatilité".



(b) ACF mené sur les valeurs absolues des rentabilités *tick-by-tick* (action Analog Devices durant la journée du 05/02/2001). On constate une autocorrélation significative entre ces valeurs absolues.



(c) ACF mené sur les valeurs absolues des rentabilités journalières de l'action Analog Devices durant la période du 01/01/1994 au 01/01/2008.



(d) ACF mené sur les valeurs absolues des rentabilités hebdomadaires de l'action Analog Devices durant la période du 01/01/1994 au 01/01/2008.

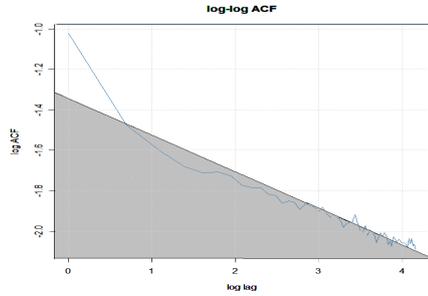
Fig. 2.4: Présence des bouffées de volatilité dans les données *tick-by-tick*, journalières et hebdomadaires.

Pour quantifier le phénomène des bouffées de volatilité, Tsay (2002) et Bouchaud (2002) remarquent que l'ACF des rentabilités absolues descend de manière hyperbolique. Autrement dit, avec l'axe Y représentant le logarithme de l'ACF et l'axe X celui du retard considéré ("*lag*"), une ligne droite peut être tracée entre Y et X comme illustré dans la figure 2.5 page 81. La pente de cette droite correspond au paramètre β de mémoire longue présenté dans la section 2.1.1.

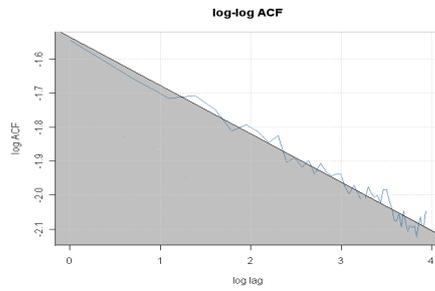
Nous pouvons ainsi estimer β pour mesurer la mémoire d'une série de rentabi-

2.1. Propriétés statistiques des séries de rentabilités

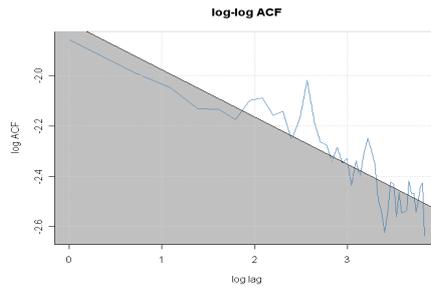
lités absolues. Cette démarche est illustrée dans la figure 2.5 avec les rentabilités absolues de l'action Analog Devices Inc. (ADI) échangée sur le NYSE.



(a) Loglog plot de l'ACF des rentabilités absolues. Fréquence : *tick-by-tick*. La pente de la ligne bleue est estimée à -0.181 .



(b) Loglog plot de l'ACF des rentabilités absolues. Fréquence : journalière. La pente de la ligne bleue est estimée à -0.14 .



(c) Loglog plot de l'ACF des rentabilités absolues. Fréquence : hebdomadaire. La pente de la ligne bleue est estimée à -0.19 .

Fig. 2.5: Estimation du paramètre β . Données utilisées : les rentabilités *tick-by-tick*, journalières et hebdomadaires de l'action ADI.

En constatant que $0 < \beta < 1$, on conclut à la présence de mémoire longue dans les rentabilités absolues d'ADI. Pour généraliser cette conclusion, nous avons estimé le β pour chaque série présentée dans la section 2.1.1 : les résultats exposés dans le tableau 2.5 confirment la présence de mémoire longue dans les rentabilités absolues : β est compris entre $(0, 1)$ pour presque toutes les séries étudiées.

Ce résultat est cohérent avec les constats de Ding, Granger, et Engle (1993) et Granger et Hyung (2004) qui soutiennent la présence de mémoire longue dans la volatilité.

Remarquons également que la fiabilité de cette conclusion diminue avec la baisse de la fréquence d'observation. Cela confirme encore une fois que la distri-

2.1. Propriétés statistiques des séries de rentabilités

Tab. 2.5: Estimation de β pour les rentabilités absolues de trois fréquences différentes.

Numéro	Titre	<i>tick-by-tick</i>	Journalière	Hebdomadaire
1	AA	-0.1833	-0.1691	-0.1209
2	AXP	-0.1842	-0.1954	-0.0483
3	BA	-0.1842	-0.1636	-0.0763
4	BAC	-0.1845	-0.1391	-0.1289
5	CAT	-0.1854	-0.1194	-0.1673
6	CSCO	-0.1854	-0.1194	-0.1673
7	CVX	-0.1854	-0.1194	-0.1673
8	DD	-0.1854	-0.1194	-0.1673
9	DIS	-0.1857	-0.1505	-0.3047
10	GE	-0.1876	-0.0985	-0.1081
11	HD	-0.1908	-0.1618	0.0440
12	HPQ	NA	-0.2186	0.0026
13	IBM	-0.1909	-0.1286	-0.2452
14	INTC	-0.1871	-0.2837	-0.1617
15	JNJ	-0.1906	-0.1860	-0.2969
16	JPM	-0.1911	-0.1439	-0.1885
17	KFT	NA	-0,1681	0,2094
18	KO	-0.1911	-0.0915	-0.2194
19	MCD	-0.1919	-0.2075	-0.2482
20	MMM	-0.1899	-0.0677	-0.0606
21	MRK	-0.1922	-0.2477	-0.0751
22	MSFT	-0.1934	-0.0618	-0.0488
23	PFE	-0.1959	-0.2131	0.0000
24	PG	-0.1958	-0.0682	-0.1963
25	T	-0.1781	-0.1250	-0.1351
26	TRV	NA	-0.1340	-0.1219
27	UTX	-0.1808	-0.1914	-0.0947
28	VZ	-0.1809	-0.1462	-0.0862
29	WMT	-0.1810	-0.2451	0.1458
30	XOM	-0.1812	-0.2369	-0.2865
Moyenne		-0.187	-0.144	-0.128
Écart-type		0.0056	0.057	0.093

On constate que les β des rentabilités *tick-by-tick* sont relativement stables. (Ils sont compris entre -0.178 et -0.196 avec une moyenne de -0.187 et un écart-type de 0.0056). Cette stabilité diminue avec la baisse de la fréquence d'observation, le constat de mémoire longue devient donc moins fiable. Ceci montre encore une fois que les rentabilités financières tendent vers une distribution *i.i.d.N* avec l'augmentation de l'intervalle d'observation.

bution des rentabilités financières ressemble de plus en plus à une loi normale quand l'intervalle d'observation augmente, ce qui constitue un résultat classique. Ce phénomène, nommé "Aggregational Gaussianity", est étudié en détail dans le paragraphe suivant.

III. Phénomènes "multi-échelle"

La notion d'auto-similarité est introduite en finance par Mandelbrot (1963) pour décrire les séries de rentabilités dont la distribution reste stable quelle que soit l'échelle de temps considérée.

Dans un processus auto-similaire, les rentabilités intra-day, journalières et hebdomadaires doivent partager la même forme de distribution. Ce phénomène est défini par l'équation (2.18).

$$\{X(ct_1), \dots, X(ct_k)\} \sim \{c^H X(t_1), \dots, c^H X(t_k)\} \quad (2.18)$$

où $H > 0$ et $c, k, t_1, t_2, \dots, t_k \geq 0$ sont des paramètres donnés.

Le paramètre H est souvent nommée "indice d'auto-similarité" ou "exposant d'échelle".

Par exemple, soit $X(t_1), \dots, X(t_k)$ une suite de rentabilités journalières, alors, en imposant $c = 5$, la suite $X(5t_1), \dots, X(5t_k)$ représente les rentabilités hebdomadaires¹². L'équation (2.18) implique que :

1. les rentabilités hebdomadaires sont approximativement égales aux rentabilités journalières multipliées par 5^H .
2. quelle que soit la valeur de c , cette même relation est vérifiée sans exception. Par exemple, les rentabilités mensuelles doivent correspondre aux rentabilités journalières fois 20^H ¹³.

Les processus *i.i.d.N* sont auto-similaires avec $H = 0.5$. Et ceux engendrés par la loi Pareto-Levy stable le sont avec $H \in (0.5, 1)$.

En revanche, pour certains processus, notamment celui suivi par des séries des rentabilités financières, l'indice H dépend de l'échelle du temps considérée (cf. le paramètre c). Dans ce cas de figure, on parle du phénomène "multi-échelle" (Calvet et Fisher 2002, Matteo 2007).

Le coefficient de Hurst généralisé (CHG) (Barabasi et Vicsek 1991, Mandelbrot 1997) est souvent utilisé pour illustrer le phénomène "multi-échelle".

12. En admettant qu'il y a 5 jours de transactions dans une semaine.

13. En supposant qu'il existe 20 jours de transactions par mois

2.1. Propriétés statistiques des séries de rentabilités

Soit $X(t)$ une série temporelle de longueur $n \in N$, alors, on peut calculer le montant k_q selon l'équation suivante :

$$k_q(\Delta t) = \frac{\sum_{t=1}^{(n-\Delta t)} |X(t + \Delta t) - X(t)|^q}{(n - \Delta t)|X(t)|^q} \quad (2.19)$$

Où Δt est un paramètre compris entre 1 et n , q un paramètre positif, et $|X(t)|$ désigne la valeur absolue de $X(t)$.

Si $k_q(\Delta t)$ ainsi défini respecte la relation 2.20, nous pouvons calculer son CHG comme suit :

$$k_q(\Delta t) \sim \Delta t^{qH(q)} \quad (2.20)$$

Remarquons que ici, $H(1)$ coïncide avec le coefficient de Hurst présenté dans la section 2.1.1.

Ainsi, en traçant la droite $\log(k_q) \sim \log(\Delta t)$, la pente de la droite obtenue correspond au paramètre $qH(q)$ (cf. figure 2.6 page 84).

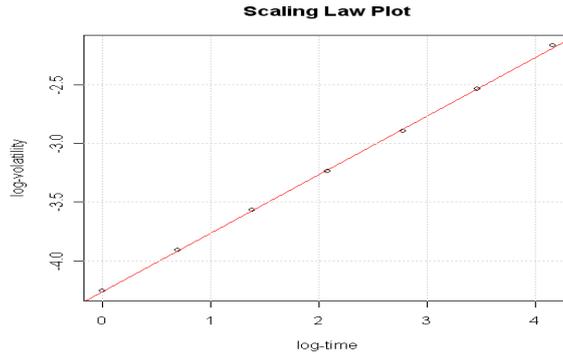


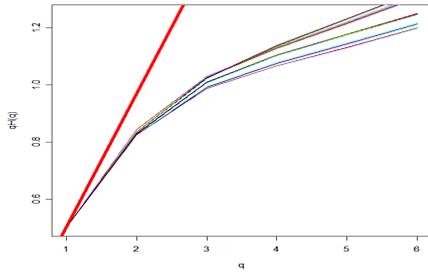
Fig. 2.6: Relation $\log(k_q) \sim \log(\Delta t)$ obtenue avec une distribution normale. Ici, la pente de la ligne droite est estimée à 0.501, elle reste stable dans le temps, car les processus *i.i.d.N* est autosimilaire avec $H = 0,5$.

La relation $qH(q) \sim q$ doit être une ligne droite si $H(q)$ est stable par rapport à q . Dans le cas contraire, on observerait un phénomène de “multi-échelle”.

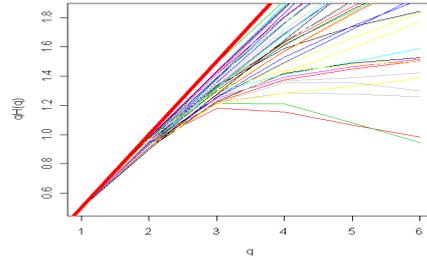
Pour vérifier l'autosimilarité des séries financières, la relation $qH(q) \sim q$ est estimée pour chaque série présentée dans la section 2.1.1. La figure 2.7 page 85 expose les résultats obtenus.

Ici, on constate qu'au lieu de retrouver une ligne droite avec un $H(q)$ unique et stable comme dans le cas d'une distribution normale, les courbes obtenues penchent vers le bas. Nous validons donc le phénomène “multi-échelle” ou “multi-fractalité” des séries financières.

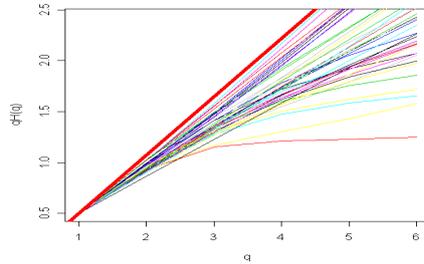
2.1. Propriétés statistiques des séries de rentabilités



(a) Fréquence : *tick-by-tick*. Aucune des séries étudiées ne dispose de CHG constant. $H(q)$ diminue avec q , car toutes les courbes colorées penchent vers le bas.



(b) Fréquence : journalière. Bien que les pentes des courbes varient considérablement dans cette figure, la distance entre la ligne rouge et les courbes colorées reste remarquable. Cela indique la non stabilité de $H(q)$.



(c) Fréquence : hebdomadaire. La distance entre la ligne rouge et les courbes colorées est flagrante, ce qui indique la non stabilité de $H(q)$.

Fig. 2.7: Relation $qH(q) \sim q$ estimée pour chacune des séries présentées dans la section 2.1.1. Ici, chaque courbe colorée représente une action, la ligne droite en rouge est ce que nous devrions observer si les rentabilités étaient *i.i.d.N*.

2.1.3 Propriétés des rentabilités financières : un récapitulatif

Dans cette section, nous avons présenté les propriétés statistiques des rentabilités financières. Nous distinguons deux types de régularités selon leur implications par rapport à la possibilité de “*market timing*”.

1. les régularités dans la moyenne.
2. les propriétés sur les moments d'ordre supérieur à 2.

En ce qui concerne la moyenne, 4 conclusions sont mises en avant :

1. Les rentabilités intra-journalières sont fortement autocorrélées. Ce phénomène souvent expliqué par la fourchette de cotation est considéré inexploitable par les investisseurs, surtout en présence des coûts de transactions.
2. Les rentabilités journalières et hebdomadaires ne sont pas autocorrélées dans le temps, ce qui confirme l'imprévisibilité des rentabilités futures à cette échelle de temps.
3. La mémoire longue n'est pas observée de manière systématique dans les innovations de prix journalières et hebdomadaires.
4. L'hypothèse *i.i.d* étant rejetée par le test BDS, on soutient une dépendance non linéaire des rentabilités, probablement très difficile à identifier précisément.

Or, si la moyenne des rentabilités s'est avérée quasi-aléatoire, leurs moments d'ordre supérieur à 2 disposent des régularités flagrantes :

1. Les rentabilités financières ont des queues épaisses, leur indice de Pareto varie entre 2 et 7.
2. Il existe une mémoire longue dans les rentabilités absolues.
3. La distribution des rentabilités tend vers une loi normale quand l'intervalle de calcul augmente. Plus précisément, le CHG des séries financières n'est pas stable.

L'implication de ce phénomène sur la modélisation des rentabilités varie selon la notion de temps adoptée :

- En adoptant une notion de temps continu, les rentabilités à toutes les fréquences doivent être prises en compte par un seul modèle. Le phénomène “multi-échelle” est donc une contrainte de modélisation à prendre en compte. Matteo (2007) fournit une discussion très détaillée à ce sujet.
- En revanche, dans le cas du temps discret, les données de différentes fréquences peuvent être caractérisées par des modèles distincts. L'impact du phénomène “multi-échelle” serait donc minimisé sur la modélisation des dynamiques financières.

À partir de ces propriétés statistiques, les économètres et les écono-physiciens ont proposé de nombreux modèles pour caractériser les séries financières. Par

exemple, on peut citer les modèles de famille GARCH (Engle 1982) pour les régularités dans la volatilité, les modèles de type STAR (Chan et Tong 1986) pour la non-linéarité de la moyenne.

Dans la prochaine section, nous allons rappeler les principaux modèles statistiques en finance en insistant sur leurs relations avec l'imprévisibilité des rentabilités futures.

2.2 De l'approche économétrique à l'impossibilité de "market timing"

Les études empiriques effectuées dans la section précédente soutiennent la conclusion suivante : si l'hypothèse d'autocorrélation semble bien écartée¹⁴, la présence d'une dépendance non linéaire s'est avérée incontestable dans les dynamiques financières.

Étant donnée l'imprévisibilité des systèmes chaotiques (défendu dans la section 2.1.1), dans cette section, nous allons nous concentrer sur les modèles stochastiques susceptibles de caractériser les séries des rentabilités successives.

À cet effet, les principaux modèles disponibles en finance sont présentés pour esquisser une image globale de la littérature en modélisation financière. Ensuite, la contribution de ces méthodes économétriques à la prévision des rentabilités sera discutée en détail pour défendre l'impossibilité de "market timing".

2.2.1 Les modèles de la famille ARMA

Les modèles de la famille ARMA (autorégressive et moyenne mobile) répondent à la formulation générale suivante :

$$\left(1 - \sum_{i=1}^p \varphi_i L^i\right) \Delta^d X_i = \left(1 + \sum_{i=1}^q \theta_i L^i\right) \varepsilon_i \quad (2.21)$$

où p, q sont les paramètres positifs nommés "ordres" du modèle. φ_i, θ_i sont les paramètres de régression. L^i est un opérateur de retard satisfaisant $L^t X_i = X_{i-t}$. Δ^d est l'ordre de différenciation. Quand $d \in]0, 1[$, l'équation (2.21) devient un modèle ARFIMA ("*Fractally Integrated ARMA*") (Baillie 1996).

14. Sauf dans le cas des rentabilités *tick-by-tick* dont l'autocorrélation est peu exploitable en présence des coûts de transaction.

En rejetant l'hypothèse d'autocorrélation, les modèles ARMA traditionnels (cf. $d = 0$) ne devraient pas convenir aux séries financières. Or, une autocorrélation non significative aux yeux des tests statistiques peut rendre valable un modèle ARMA, car les niveaux de signification exigés par ces deux critères ne sont pas identiques. Les rentabilités successives semblent se trouver dans cette situation : leur autocorrélation n'est pas validée par les tests statistiques mais celle-ci permet d'utiliser les modèles de la famille ARMA.

Par exemple, Mignon (1998) retient le modèle autorégressif pour 5 des 7 indices boursiers étudiés à une fréquence journalière ¹⁵. Tsay (2005) utilise un modèle autorégressif d'ordre 3 pour modéliser les variations hebdomadaires de l'indice CRSP.

De la même manière, le rejet de mémoire longue devrait aller de pair avec l'impossibilité d'utiliser les modèles ARFIMA. Cependant, sans que cela devienne un "fait stylisé", certaines séries de rentabilités affichent un ordre de différenciation $0 < d < 1$: celles-ci peuvent donc être caractérisées par un modèle ARFIMA.

Parmi d'autres, Mignon (1998) a conclu à la supériorité du modèle ARFIMA par rapport à celui de marche au hasard. En examinant aussi bien les rentabilités réelles que celles engendrées par des simulations de Monte-Carlo, Bhardwaj et Swanson (2006) soutiennent l'utilité du modèle ARFIMA en matière de prédiction à long terme.

2.2.2 ARCH, GARCH, et FIGARCH

Le modèle GARCH (General Autoregressive Conditional Heteroskedasticity) a la forme générale suivante :

$$X_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i X_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad (2.22)$$

où X_t désigne la t ème rentabilité observée, ε_t est *i.i.d* d'espérance nulle. $\alpha_0 > 0$, $\alpha_i, \beta_j \geq 0$, et $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$ ¹⁶.

Ce modèle, proposé par Bollerslev (1986) pour assouplir les contraintes techniques subies par le modèle ARCH (Engle 1982), caractérise avant tout la volati-

15. Les 7 indices sont respectivement : *SP500*, FT500, SBF250, *TOPIX*, TSE300, FAZ, BCI. Les deux indices mis en *italique* ne peuvent pas être prise en compte par les modèles autorégressifs.

16. Cette dernière condition peut garantir l'existence de la variance de X_t .

lité des séries de rentabilités¹⁷.

Les transformations du modèle GARCH sont nombreuses. Par exemple, Nelson (1991) utilise un "Exponential GARCH" (EGARCH) pour tenir compte de l'asymétrie entre les rentabilités négatives et positives. En rajoutant un bruit blanc dans le terme " σ_t^2 " de l'équation (2.22), Taylor (1994) propose un "stochastic volatility model"¹⁸. En introduisant un ordre de différenciation $0 < d < 0,5$ dans le modèle de volatilité stochastique, Bollerslev et Jubinski (1999) et Ray et Tsay (2000) obtiennent le "long memory stochastic volatility model"¹⁹. Ces derniers auteurs ont d'ailleurs conclu à un ordre de différenciation $d = 0,38$ en étudiant les valeurs absolues des rentabilités journalières.

2.2.3 Modèles multi-états

Si les modèles de type "GARCH" prennent en compte la non-linéarité de la variance, les modèles multi-états caractérisent celle de la moyenne. Ces modèles, spécifiés par la coexistence de plusieurs régimes d'évolution, ont retenu l'attention des chercheurs en finance depuis la crise de 1987.

Trois catégories de modèles multi-états semblent dominer la littérature en finance quantitative.

I. "Threshold models", ou modèle à seuil

Les modèles de la famille "threshold autoregressive" (désormais TAR), initialement proposés par Petrucci et Woolford (1984), sont souvent utilisés pour introduire des changements de régimes dans les séries financières.

Par exemple, le modèle TAR à deux régimes défini par l'équation (2.25),

17. Remarquons qu'en imposant $s = 0$, l'équation (2.22) devient le modèle ARCH de Engle (1982), une forme simplifiée du modèle GARCH.

18. En effet, la formulation générale d'un "stochastic volatility model" est :

$$X_t = \sigma_t \varepsilon_t, \quad (1 - \alpha_1 L - \dots - \alpha_m L^m) \ln(\sigma_t^2) = \alpha_0 + v_t \quad (2.23)$$

où L est toujours l'opérateur de retard, v_t est un *i.i.d* d'espérance nulle.

19. Le modèle LMSV est souvent formulé comme suit :

$$X_t = \sigma_t \varepsilon_t, \quad \sigma_t = \sigma \exp(u_t/2), \quad \Delta^d u_t = \eta_t \quad (2.24)$$

où $\theta > 0$, ε_t, η_t sont *i.i.d.N* d'espérance nulle, Δ^d désigne l'opération de différenciation et $0 < d < 0,5$.

peut engendrer des séquences chronologiques semblables à celles simulées dans la figure 2.8. On constate dans cette simulation des pics de volatilité (cf. 155ième observation) qui pourraient tenir compte des variations excessives, telles que le lundi noir en 1987.

$$x_t = \begin{cases} -1,5x_{t-1} + a_t & \text{si } x_{t-1} < 0, \\ 0,5x_{t-1} + a_t & \text{si } x_{t-1} \geq 0. \end{cases} \quad (2.25)$$

où a_t est *i.i.d.* $N(0, 1)$. En introduisant une fonction de transition lisse et continue,

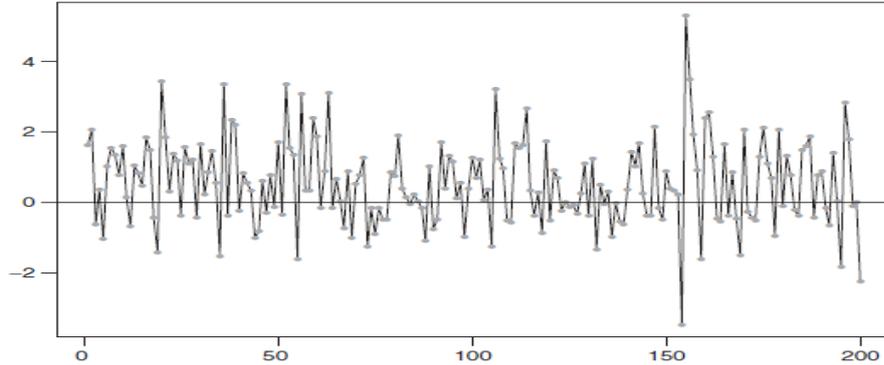


Fig. 2.8: Simulation du modèle TAR(1).

souvent supposée logistique ou exponentielle, Chan et Tong (1986) proposent le modèle STAR (Smooth Threshold Autoregressive Model) dont l'application porte souvent sur la modélisation des facteurs économiques (ex. : le taux de chômage, la variation de dividende ou la croissance du PIB). Les dynamiques boursières sont ensuite caractérisées selon l'évolution de ces facteurs fondamentaux. Par exemple, Sarantis (2001) utilise le modèle STAR pour modéliser les indices boursiers des principaux pays industriels de 1969 à 1999. En étudiant les données plus récentes (1988-2002), Jawadi et Koubaa (2007) caractérisent les indices des pays du G7 avec ce même modèle.

II. "Markov Switching Autoregressive Model"

Selon Hamilton (1989), une série chronologique x_t suit un modèle de type "Markov Switching Autoregressive" (désormais MSA), si elle satisfait :

$$x_t = \begin{cases} c_1 + \sum_{i=1}^p \phi_{1,i} x_{t-i} + a_{1t} & \text{si } s_t = 1, \\ c_2 + \sum_{i=1}^p \phi_{2,i} x_{t-i} + a_{2t} & \text{si } s_t = 2, \end{cases} \quad (2.26)$$

où a_{1t} , a_{2t} sont *i.i.d.* d'espérance nulle. s_t désigne l'état du MSA dont les probabilités de transition sont respectivement :

$$P(s_t = 2 | s_{t-1} = 1) = \omega_1, \quad P(s_t = 1 | s_{t-1} = 2) = \omega_2 \quad (2.27)$$

Si l'état futur d'un modèle TAR est calculé par sa fonction de transition, celui d'un MSA dépend des variables stochastiques dont la relation avec les données historiques n'est pas directement observable. Quand les observations historiques sont suffisamment nombreuses, certaines méthodes non-paramétriques permettent d'estimer l'état futur du modèle MSA²⁰. Néanmoins, à cause de ces approches complexes, la fiabilité de ses prédictions est sensiblement réduite par rapport aux modèles précédents.

III. Modèles d'apprentissage neuronal

Les modèles de neuro-apprentissage ("*Neural networks*") sont des associeurs universels qui font partie des techniques d'intelligence artificielle. Pour modéliser le processus d'apprentissage d'un cerveau humain, ils sont organisés en plusieurs couches ("*layers*") : la couche inférieure ("*input layer*") stocke toutes les observations historiques, la couche extérieure ("*output layer*") donne la prédiction finale. Et entre ces deux, on trouve un certain nombre de couches cachées ("*hidden layers*"), dont chacune contient un nombre fixe de noeuds intermédiaires. La figure 2.9 illustre la forme d'un modèle neuro-apprentissage (désormais NA) avec 1 seule couche cachée contenant deux noeuds intermédiaires.

L'entrée totale du noeud " d_j " est donnée par l'équation (2.28)

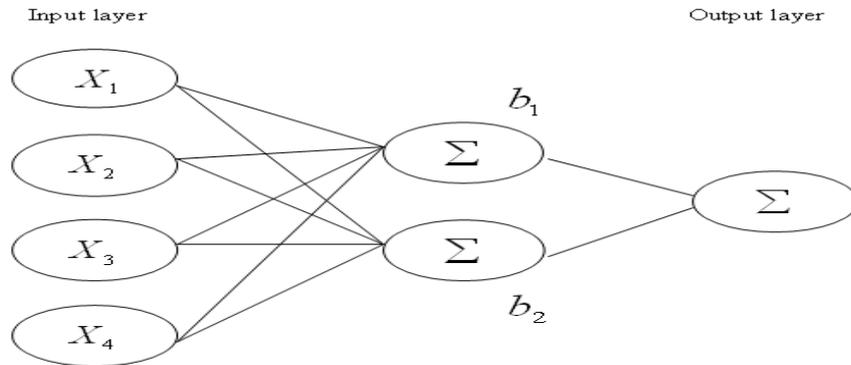


Fig. 2.9: Présentation d'un modèle NA à une couche cachée et deux noeuds intermédiaires.

$$b_j = f\left(\sum_{i=1}^n X_i w_{i,j} - w_{0,j} \varphi_j\right) = f(A_j) \quad (2.28)$$

où $w_{i,j}$ désigne le poids de X_i pour le noeud d_j , φ_j le biais de l'entrée X_i , $w_{0,j}$ le poids du biais correspondant, et $f(A_j) = \sum_{i=1}^n X_i w_{i,j} - w_{0,j} \varphi_j$. Généralement, $f(A_j)$ est supposé être une fonction logistique du vecteur A_j .

20. Pour plus d'informations à ce sujet, voir le chapitre 22 du livre Hamilton (1994).

Au départ, les paramètres du modèle sont fixés de manière aléatoire, le processus de "training" s'arrête quand les racines carrées des résidus (l'indicateur "RMSE" qui mesure la distance entre les données observées et les prédictions du modèle) atteignent son minimum.

Comme Zhang l'a résumé, "*neural networks are data-driven self-adaptive methods in that there are few a priori assumptions about the models under study* (Zhang, Eddy Patuwo, et Y. Hu 1998, p.35)." Cette approche, dont la forme peut s'adapter à chaque échantillon analysé, offre souvent des meilleures caractérisations que les modèles traditionnels.

En revanche, cet avantage est aussi son principal défaut : le modèle obtenu n'étant soutenu par aucune théorie économique, le risque d'*overfitting* est plus élevé pour les modèles MSA que pour les modèles traditionnels. Les prédictions fondées sur les "neural networks" ne sont pas toujours de meilleure qualité que les modèles linéaires. Pour mieux comprendre la nature de cette question, l'effet des approches économétriques sur la possibilité de "market timing" est discutée en détail dans la prochaine section.

2.2.4 Retour sur l'impossibilité de "market timing"

Certains modèles économétriques caractérisent les séries financières en mettant en évidence des paramètres significatifs issus de l'estimation statistique. Cependant, comme Fama (1970) l'a précisé, la signification statistique n'implique pas toujours une signification financière²¹. L'usage des modèles économétriques pour faire du "market timing" a donc suscité de nombreux tests empiriques.

La méthode générale suivie par ces tests consiste à simuler une stratégie qui vend l'actif quand le modèle étudié prévoit une baisse et achète dans le cas contraire. La différence de gains entre la stratégie simulée et la stratégie "buy and hold" indique l'effet du modèle sur la possibilité de "market timing".

Parmi les modèles présentés précédemment, ceux de la famille GARCH sont fondés sur l'autocorrélation de la volatilité, ils ne peuvent donc pas prévoir le signe des variations futures. L'impact de ces modèles doit être neutre sur la possibilité de "market timing".

La dynamique chaotique dont l'imprévisibilité est défendue dans la section 2.1.1 (page 70) est également compatible avec l'impossibilité de battre le marché.

21. Même si les modèles sont statistiquement valides, leur exploitation pratique reste sujette à caution.

En revanche, les modèles ARMA, ARFIMA, STAR, MSA et NA sont susceptibles de prédire le sens des variations futures. Les travaux de simulation fondés sur ces modèles sont étudiés dans cette section pour défendre le point de vue suivant : bien que la dynamique boursière puisse être caractérisée par les modèles économétriques, les prévisions fournies par ces modèles ne permettent pas de battre la "buy and hold".

Comparés avec celle des modèles non-linéaires, l'estimation des modèles ARMA introduit moins d'éléments incertains, les paramètres obtenus sont souvent plus fiables. En revanche, comme ces modèles fournissent des prédictions linéaires, leur variation est souvent très limitée. Le sens des rentabilités futures demeure donc difficilement déductible. Comme illustré dans la figure 2.10 (page 2.10), les prédictions fournies par le modèles AR(3) ne varient que légèrement dans le temps, celle-ci converge d'ailleurs rapidement vers la moyenne de l'échantillon. Par conséquent, anticiper le sens des rentabilités futures demeure difficile malgré la signification statistique du modèle estimé. Cette analyse rejoint la conclusion

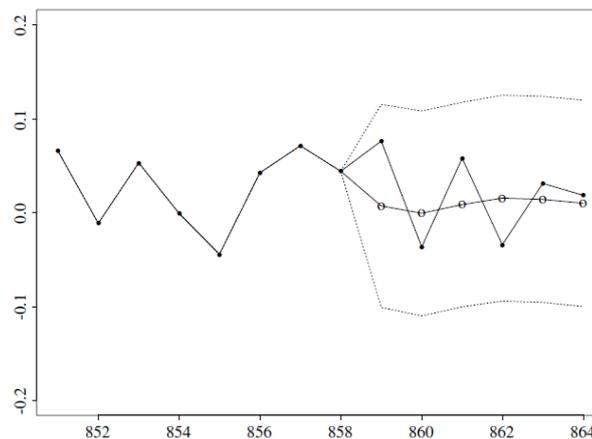


Fig. 2.10: Prédiction de l'indice CRSP pondéré avec un modèle AR(3).

Données utilisées pour la régression : l'indice CRSP pondéré observé pendant la période 1926-1997 à la fréquence mensuelle. Les lignes discontinues forment l'intervalle de confiance à 5%. Les points noirs présentent les prédictions à 1-6 pas. La courbe continue désigne les observations réelles. La prédiction commence par la 858ème observation. Source : reproduite à partir de la figure 2.5 de Tsay (2002).

de Lane, Peel, et Raeburn (1996) qui, après avoir testé le pouvoir de prévision de plusieurs modèles linéaires et non-linéaires, ont conclu que "les prévisions obtenues affichent tellement peu de variation que la possibilité de formuler des stratégies exploitables doit être remise en question (Lane, Peel, et Raeburn 1996, p.424)."

Les modèles non linéaires, sensés reproduire les valeurs extrêmes des rentabilités, ne peuvent pas non plus augmenter la profitabilité des prévisions économé-

triques.

Selon Bellgard et Goldschmidt (1999), pour réussir une stratégie de type "market timing", il faut prévoir le sens des variations futures. Les indicateurs traditionnels de la qualité des modèles statistiques²² ne sont pas toujours appropriés quand il s'agit d'estimer la possibilité de battre la "buy and hold".

Certaines études, comme Mignon (1998) et Bhardwaj et Swanson (2006), qui utilisent ces indicateurs pour comparer des modèles statistiques, ne délivrent pas toujours d'informations propices à l'égard de la possibilité de "market timing".

Grabbe (1996) propose d'évaluer cette possibilité avec le taux de réussite obtenu en prévoyant le sens de variations futures, ou le "hit-score (Hellström et Holmström 1998)". Selon Grabbe, le "hit score" d'un modèle doit être supérieur à 60% pour concevoir une stratégie profitable de type "market timing".

Suivant cette logique, Sarantis (2001) teste la capacité des modèles STAR à prédire le sens des variations futures. Bien que cet auteur ait démontré la signification statistique des modèles utilisés, leur pouvoir de prédiction s'est avéré peu satisfaisant : le meilleur taux de réussite est autour de 60% pour les indices boursiers anglais, français et canadien. Ce taux est estimé à 50% pour l'indice américain, et à 45% pour le marché japonais. Ces observations soutiennent sans doute l'impossibilité de battre la "buy and hold".

Dans le premier chapitre du livre de Dunis et Williams (2003), les auteurs ont simulé les performances des modèles ARMA, moyennes mobiles, logit et neuro-apprentissage sur les taux de change EURO/USD du 19/05/2000 au 03/07/2001 (observés à la fréquence journalière). Bien que le modèle de NA semble plus performant que ses concurrents, son taux de réussite en sens de variation n'est que de 57,24%. Ces auteurs ont ainsi conclu que les méthodes quantitatives sont plus appropriées dans un cadre du "stock picking" que du "market-timing".

Cette conclusion est aussi confirmée par les tests statistiques. En appliquant un test non-paramétrique, Diebold et Nason (1990) montrent que les modèles non-linéaires n'engendrent pas de meilleures prévisions que les linéaires. En testant les variations journalières des prix des contrats futurs sur les indices SP500, Nikkei, DAX, FTSE et CAC40, Hsieh (1995) conclut que "la volatilité future est prévisible, mais pas le sens des variations (Hsieh 1995, p.55)." En analysant les séries de taux de change entre la livre sterling et le dollar américain, Zhang et Berardi (2001) démontrent l'équivalence des modèles de "neuro-apprentissage" et de "marche au hasard" en terme de prédiction. Ce résultat est confirmé par Tsay (2005) qui, en

22. Par exemple, RMSE, AIC (Akaike 1973), SIC (Schwarz 1978), etc.

étudiant les variations mensuelles de l'action IBM, n'a pas constaté de différence significative entre les modèles linéaires et non-linéaires.

Les analyses dans cette section convergent vers une conclusion unique : bien que certains modèles économétriques soient significatifs au regard des séries financières, les régularités constatées dans ces rentabilités restent compatibles avec l'impossibilité de "*market timing*".

Conclusion

Ce chapitre propose une étude générale des propriétés statistiques des séries de rentabilités. L'objectif d'une telle étude était de montrer l'impossibilité de "*market timing*" sur ces chronologiques pourtant structurées autour de régularités importantes. La figure 2.11 récapitule l'ensemble des possibilités offertes par les différents modèles disponibles à cet effet.

Exceptées les données *tick-by-tick* qui sont négativement autocorrélées à cause de la fourchette de cotation, les séries des rentabilités journalières et hebdomadaires sont stationnaires à l'égard des tests effectués. Les tests d'autocorrélation et de mémoire longue rejettent également la présence des dépendances linéaires dans ces séries, ce qui conduit à la réalisation des tests BDS. Le résultat de ces derniers tests semble concluant : les séries financières ne sont pas *i.i.d.*, il y existerait probablement une dépendance non linéaire qui les structurerait.

On distingue deux types de dépendances non linéaires :

- dépendances dans la volatilité (i.e. faits stylisés), souvent prises en compte par les modèles de la famille GARCH.
- dépendances dans la moyenne, généralement saisies par les modèles multi-états.

Après ces caractérisations statistiques, l'impact des régularités constatées sur la possibilité de "*market timing*" est discuté en 3 étapes :

1. Bien que l'autocorrélation des séries financières soit rejetée par les tests statistiques, celle-ci est parfois suffisamment présente pour rendre significatif les modèles autorégressifs. Or, les modèles ARMA obtenus délivrent peu d'information sur le sens des variations futures, ils sont donc compatibles avec l'impossibilité de "market timing".
2. Par définition, les dépendances dans la volatilité n'aident pas à prédire le sens des rentabilités futures. Malgré leur contribution à l'évaluation des risques financiers, ces régularités sont peu exploitables par les stratégies de

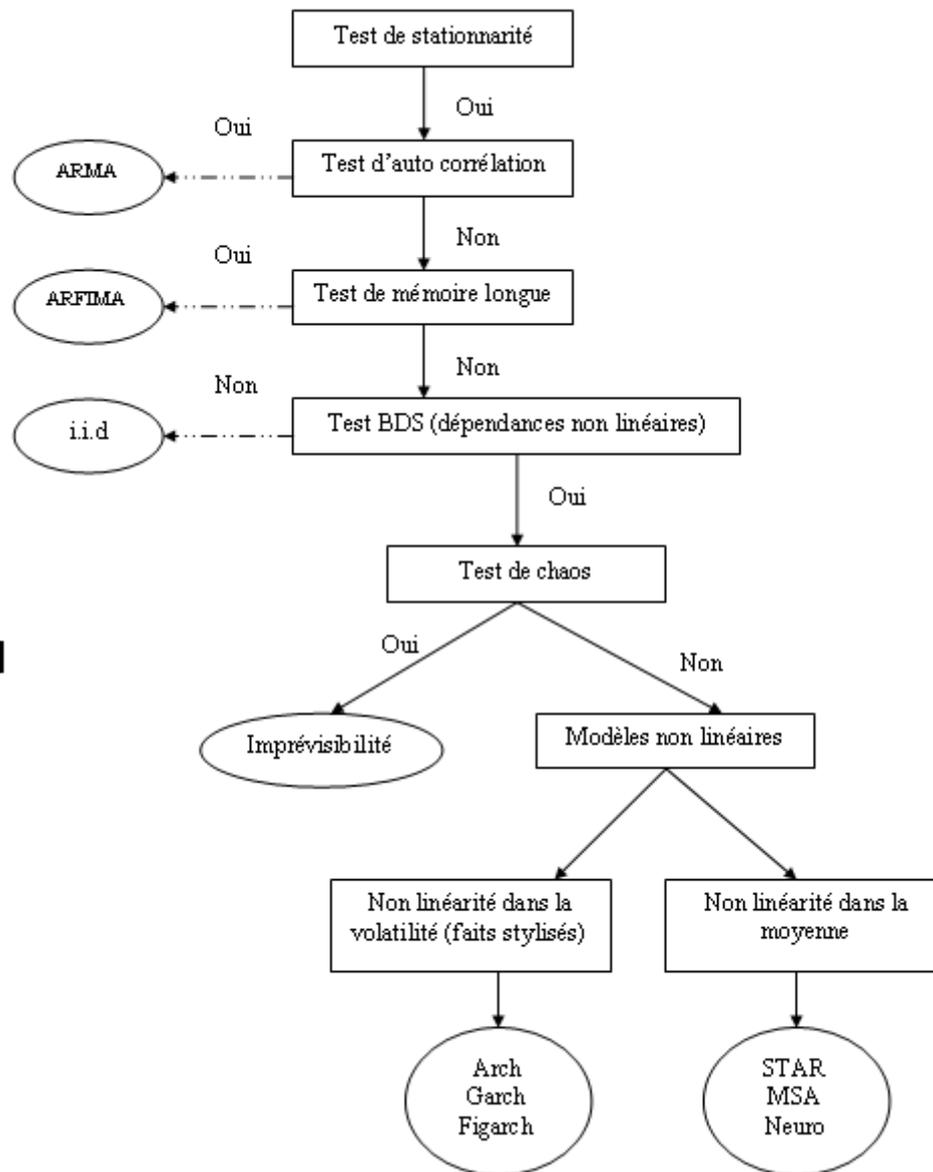


Fig. 2.11: Caractérisation des séries de rentabilités.

type “*market timing*”.

3. Les dépendances non linéaires dans la moyenne sont, à priori, incompatibles avec l'impossibilité de “*market timing*”. Cependant, les modèles multi-états fondés sur ces régularités nécessitent souvent un grand nombre de paramètres à estimer. La fiabilité de leurs prédictions est donc considérablement pénalisée par ces difficultés d'estimation. De nombreux travaux empiriques ont conclu à l'équivalence des modèles linéaires et non linéaires en matière de prédiction.

Bien entendu, la relation entre les méthodes économétriques et l'impossibilité de “*market timing*” pourrait être modifiée avec le développement des techniques de modélisation (Lo et MacKinlay 2002).

En revanche, l'avantage des modèles innovants ne dure qu'un temps très limité, car ces techniques seront toujours découvertes par d'autres investisseurs, ce qui diminuera considérablement leur pouvoir de prévision. D'après Derman (2002), la durée de vie d'un nouveau modèle quantitatif ne dépasse pas 3 mois. Les investisseurs concurrents trouvent la même technique à une vitesse absolument stupéfiante.

Pour conclure, l'impossibilité de “*market timing*” est non seulement compatible avec les principaux courants théoriques en finance, mais aussi soutenue par les faits empiriques observés dans les données historiques.

3. COMPLEXITÉ DE KOLMOGOROV : UNE APPLICATION DE LA THÉORIE ALGORITHMIQUE DE L'INFORMATION AUX SÉRIES DE RENTABILITÉS

Résumé

Dans ce chapitre, nous estimons le degré d'aléa des séries financières. Si les méthodes statistiques ont toujours dominé la saisine du hasard et de la régularité en finance, nous proposons ici de les mesurer dans un cadre différent issu de la théorie de la calculabilité, i.e., une branche en informatique fondamentale qui étudie la possibilité de résoudre un problème donné avec un ordinateur.

À cet effet, une approche générale est introduite pour estimer la “*complexité de Kolmogorov*” des séries de rentabilités : après un processus de “discrétisation”, des algorithmes de compression sans perte sont utilisés pour détecter des structures régulières qui ne sont pas toujours visibles aux yeux des tests statistiques. Cette approche est illustrée aussi bien avec des données de simulation qu'avec des séries de rentabilités financières.

En étudiant le degré d'aléa des principaux marchés internationaux (i.e., NYSE, NASDAQ, Euronext-Paris) observés à une fréquence *tick-by-tick*, on constate un degré d'aléa plus élevé pour Euronext-Paris que pour les deux marchés américains. Nous expliquons ce résultat par le fait que l'auto-corrélation des volatilités inter-journalières est plus élevée aux États-Unis qu'en France.

Introduction

Jusqu'à présent, les méthodes quantitatives que nous avons évoquées s'appuient principalement sur les inférences statistiques et la théorie des probabilités. L'orientation de la finance de marché vers cette théorie est vraisemblablement motivée par l'incertitude des agents économiques dans leurs anticipations des rentabilités futures.

Il semble que le hasard et les probabilités forment un couple inséparable qui domine la compréhension des dynamiques boursières. Cependant, l'aléa est également étudié dans la théorie de la calculabilité, une branche de l'informatique fondamentale et de la logique mathématique, qui s'interroge sur la possibilité de résoudre une question donnée à l'aide des méthodes computationnelles. Cette branche théorique n'a connu, à ce jour et à notre connaissance, que peu d'applications dans le champ de la finance¹. Elle ouvre pourtant des perspectives tant théoriques que pratiques séduisantes (Zenil et Delahaye (2011) proposent par exemple une première investigation sur le sujet).

L'objet de ce chapitre est donc de présenter la notion de hasard telle qu'elle est comprise dans la théorie de la calculabilité, ainsi que l'intérêt qu'elle représente pour les analyses des dynamiques financières. La présentation théorique qui suit (la section 3.3) est fondée sur les livres de Downey et Hirschfeldt (2010) et de Delahaye (1994) qui offrent respectivement une synthèse complète et détaillée des principaux résultats obtenus autour de la notion de hasard en calculabilité.

Une idée générale qui parcourt ce chapitre est que toute forme de régularité dans un objet (un signal, une chronique, une image, un texte) peut être exploitée pour en réduire la taille à l'aide d'algorithmes de compression. Plus ces objets sont "hasardeux", moins qu'ils peuvent être compressés.

Une particularité des travaux en informatique est que leurs objets d'analyse sont souvent des séries binaires composées de 0 et de 1, donc assez éloignées en apparence des chroniques financières. Les algorithmes de compression peuvent être directement utilisés sur ces séries binaires pour y détecter d'éventuelles régularités cachées. Or, les rentabilités financières, souvent représentées en nombres réels, ne sont pas adaptées à une telle application. Si donc elles cachent des régularités, tenter de les révéler directement à l'aide des algorithmes de compres-

1. On peut citer, par exemple, Mansilia (2004) pour une première application de la théorie de la calculabilité en finance. Comme nous allons le présenter dans la section 3.4.1 page 116, ce travail utilise l'entropie de Shannon pour estimer la complexité de Kolmogorov. Étant donnée la différence conceptuelle entre ces deux notions, l'indicateur d'aléa proposé par cet auteur pourrait encore faire l'objet d'amélioration.

sion s'avérerait fallacieusement probant, nous détaillons ce point ultérieurement. C'est pourquoi nous proposons, dans ce chapitre, une procédure de transformation nommée "discrétisation-effacement" qui rendra possible l'usage des algorithmes de compression dans les études des séries financières.

Notre approche sera illustrée par plusieurs exemples concrets, qui contiennent aussi bien des données de simulation que de vraies séries boursières observées sur les marchés financiers. À l'aide de ces expériences concrètes, nous espérons mettre en évidence l'intérêt de la démarche algorithmique dans la détection des régularités financières, ainsi que son rôle complémentaire au regard de l'approche probabiliste.

Ce chapitre est organisé en 3 parties : on présente tout d'abord la notion de hasard en calculabilité. La procédure de "discrétisation-effacement" qui transforme les données financières en séries de nombres entiers est alors détaillée. Enfin, des exemples concrets illustrent l'utilité de l'approche algorithmique dans les études des séries financières.

Pour bien illustrer des concepts en théorie de la calculabilité, certains exemples dans ce chapitre sont empruntés au manuel Wolper (1991).

3.1 Le hasard et la science, un couple d'adversaires ?

Comment l'aléa s'oppose-t-il à la régularité ? Dans quelle mesure une série de nombres réels peut-elle être considérée comme aléatoire ? Comment quantifier le degré de hasard (ou de régularité) caché derrière une telle série ? Telles sont les questions que nous allons étudier dans cette section.

En langage commun, le terme "hasard" est souvent employé pour exprimer l'ignorance des causes. L'expression "une rencontre par hasard" qualifie "le hasard", comme la cause d'éléments non maîtrisés. Autrement dit, cette expression signifie : "on ne peut pas savoir comment on s'est rencontré."

Remarquons que la notion de hasard se réfère toujours à un système de préconceptions donné. Dans l'exemple ci-dessus, ce système est fondé sur les connaissances des individus concernés : deux personnes pensent s'être rencontrées par hasard car elles ne peuvent pas l'expliquer en faisant les déductions ou inductions à partir des informations dont elles disposent. Or, s'il existait une tierce personne qui connaissait exactement les trajectoires des deux individus (par exemple un de leurs amis qui a organisé la rencontre), celle-ci considérerait leur rencontre

comme une conséquence logique et non un fruit du hasard.

De ce point de vue, **le hasard est un terme qui regroupe tous les éléments incompréhensibles dans un système de préconceptions donné.**

Il est nécessaire de préciser que la quantité d'aléa dans un système de préconceptions donné (par exemple, un système d'axiomes pour la géométrie) dépend uniquement du système, et non de la capacité déductive considérée. Parfois, un système de préconceptions est tellement compliqué qu'il est difficile de connaître toutes les propositions qui y sont déductibles. Cependant, tout élément non connu n'appartient pas au champ du hasard. Il existe une différence entre les éléments **incompréhensibles** et **non déduits**. Par exemple, un professeur en mathématiques sait déduire probablement plus de propositions dans le système Euclidien qu'un étudiant de classe terminale. Cela ne signifie pas que le hasard en géométrie Euclidienne est plus large pour l'étudiant que pour son professeur. Au contraire, en se référant au système Euclidien, ils sont exposés au même "champ de hasard". Ce qui crée la différence entre le professeur et son étudiant est seulement leur capacité de déduction.

La science qui est chargée d'améliorer notre compréhension du monde doit remplir deux missions principales : (I), en restant dans le système de connaissance actuel, accroître le nombre de déductions réalisées. (II), améliorer notre système de préconceptions.

Une question intéressante est donc de savoir si, pour chaque système de préconceptions, il existe toujours une fatalité du hasard. Autrement dit, peut-on trouver un "système complet" où toutes les propositions formulables y sont démontrables ? Cette question a été résolue par les théoriciens en calculabilité, qui ont montré que dans un système de démonstration donné², il existe toujours des propositions indémonstrables (théorème d'incomplétude de Gödel (1931)). Ces propositions impossibles à démontrer constituent donc une forme de hasard dans chaque système mathématique considéré.

Pour conclure, le hasard existe uniquement si l'on se réfère à un système de préconceptions. Pour chaque système mathématique, il existe toujours un champ de hasard qui marque la limite des déductions réalisables. Cependant, le hasard ne marque pas la limite absolue de la science, car l'objectif de cette dernière

2. Ici, le système de démonstration doit être raisonnable. Selon Delahaye (2009b), un système est raisonnable s'il satisfait les trois conditions suivantes : "d'abord, on ne souhaite pas que les preuves conduisent à démontrer des affirmations contradictoires. Ensuite, on veut que les preuves permettent de démontrer les énoncés élémentaires vrais de l'arithmétique, par exemple que 25 est un carré, ou qu'il existe une infinité de nombres premiers. Enfin, si quelqu'un utilise cette définition des preuves pour démontrer un théorème, on exige que la preuve soit vérifiable sans risque d'erreur et d'une manière automatique." (Delahaye 2009b, p.88)

n'est pas seulement de développer de nouvelles déductions, elle vise également à améliorer notre système de préconceptions, la racine de toutes les démonstrations scientifiques.

En mathématiques, le concept d'aléa est exploré dans deux branches bien distinctes :

1. la théorie des probabilités,
2. la théorie de la calculabilité.

L'objectif de cette section est de présenter en détail ces deux différentes interprétations du hasard, afin de fournir les éléments nécessaires à la compréhension de notre approche algorithmique destinée aux analyses des séries boursières.

3.2 Une approche probabiliste

Face à une situation non maîtrisée, on peut utiliser les probabilités pour modéliser des conséquences éventuelles. Par exemple, quand un chef d'entreprise prévoit un chiffre d'affaires de 1M euros avec un niveau de confiance à 80%, il décrit en réalité la conséquence éventuelle d'une situation non maîtrisée en termes de probabilités.

Étant donnée la distinction entre les éléments incompréhensibles et non maîtrisés, ici, il faut préciser que les probabilités ne modélisent pas uniquement le hasard, elles s'appliquent à toutes les situations non maîtrisées. Dans l'exemple du chiffre d'affaires, l'incertitude du chef d'entreprise peut être causée par une insuffisance de déduction et non attribuée à un hasard incompréhensible. Autrement dit, il a peut-être à sa disposition toutes les informations nécessaires au calcul de son futur chiffre d'affaires, l'incertitude peut résulter uniquement d'une incapacité de déduction.

Suivant cette logique, les situations non maîtrisées sont modélisées par des variables aléatoires. Prévoir son chiffre d'affaires avec un niveau de confiance de 80%, signifie que la conclusion du chef d'entreprise n'est pas le seul scénario du futur, les éléments imprévisibles influenceront le chiffre d'affaires, et le résultat d'une telle intervention est modélisé par une variable aléatoire dont on retient souvent l'espérance mathématique.

Ainsi, en probabilité, distinguer le hasard de la régularité revient à reconnaître les réalisations idiosyncratiques des variables aléatoires. Pour cela, la théorie propose des tests statistiques qui estiment, avec un certain niveau de confiance, si les données observées proviennent d'une variable aléatoire ayant des propriétés

de distribution spécifiques.

Par exemple, les prix successifs sont souvent supposés suivre une “marche au hasard” :

$$p_t = p_{t-1} + \varepsilon_t$$

Où ε_t désigne la rentabilité à l’instant t engendrée par une variable aléatoire *i.i.d.* $N(0, \Omega)$.

Pour vérifier la validité de cette hypothèse, nous pouvons recueillir les données réelles sur les marchés financiers, et conduire des tests statistiques qui donnent le degré de ressemblance entre l’échantillon observé et les réalisations d’une variable aléatoire *i.i.d.* N . Ce degré de ressemblance exprimé sous forme de “p-value” (définie comme le risque couru en rejetant l’hypothèse nulle), peut être considéré comme un indicateur du degré d’aléa dans les données observées.

Comme précisé précédemment, cette approche probabiliste domine incontestablement la recherche en finance et beaucoup d’autres domaines en sciences humaines. Cependant, si les probabilités sont un outil efficace pour décrire les situations non maîtrisées, la définition mathématique de celles-ci reste encore un sujet à débattre. Qu’est-ce que la probabilité ? Comment la définir ? Telles sont des questions difficiles à résoudre dans le cadre de la théorie des probabilités. Delahaye (1994) cite Zvonkin et Levin (1970) pour illustrer ce problème :

“En effet, les probabilités sont en général interprétées à l’aide de l’argument suivant : lorsqu’on effectue de nombreux tests, le rapport du nombre de cas favorables sur le nombre de cas total donnera toujours un nombre proche de la probabilité de l’événement considéré, et à la limite donnera toujours exactement cette probabilité. Cependant, dire “toujours” serait inexact : au sens strict cela se produit avec une probabilité de 1 (et dans le cas fini, avec une probabilité proche de 1). Et donc en s’y prenant ainsi, on définit la probabilité d’un événement quelconque en utilisant le concept d’un événement se produisant avec une probabilité de 1. Ceci est évidemment circulaire.”

De ce point de vue, la probabilité est plus un outil pour décrire et modéliser le hasard qu’une définition précise de celui-ci. Face à la question “qu’est-ce que le hasard ?” l’approche probabiliste semble incapable de fournir des réponses satisfaisantes.

Or, l’approche algorithmique que nous allons présenter dans la prochaine section est une théorie complète sur la notion de hasard et régularité. Cette approche propose non seulement des outils empiriques permettant de “mesurer” le degré d’aléa dans les données observées, mais aussi les définitions de hasard qui intègrent la différence entre les éléments non déduits et incompréhensibles.

3.3 Une approche algorithmique

En informatique fondamentale, la notion de hasard a été étudiée dans le cadre des suites infinies composées de 0 et de 1 (par exemple, “10011001...”), ayant pour objectif de résoudre la problématique suivante : parmi ces suites infinies, lesquelles peuvent être considérées comme véritablement aléatoires ?

Trois pistes ont été exploitées pour répondre à cette question :

1. les suites aléatoires qui réussissent tous les tests statistiques effectifs (Martin-Löf 1966),
2. les suites aléatoires qui ne sont battables par aucune stratégie de pari effective (Schnorr 1971),
3. et les suites aléatoires qui ne sont compressibles par aucun programme de compression (Kolmogorov 1965).

L’objectif des développements suivants est de présenter ces trois définitions des suites aléatoires qui ont beaucoup inspiré l’approche empirique qui sera introduite pour quantifier le degré d’aléa des séries financières. Pour que les lecteurs puissent s’adapter à la terminologie informatique, nous avons présenté dans l’annexe K quelques notions élémentaires en calculabilité qui faciliteront la compréhension de ce chapitre.

3.3.1 Hasard et tests statistiques

En théorie des probabilités, les statisticiens vérifient l’origine des données en réalisant des tests statistiques. En suivant cette logique, Martin-Löf (1966) propose de définir les suites aléatoires composées de 1 et de 0 par celles qui réussissent tous les tests statistiques effectifs³.

À chaque suite infinie s composée de 0 et de 1, nous pouvons associer un nombre réel x compris entre $[0, 1]$ dont le développement binaire est s . Par exemple, la suite $s = \underline{0110011} \dots$ est associée au nombre réel $x = 0.3984375 \dots$ ⁴.

3. La signification du terme “effectif” est présentée en détails dans l’annexe K. Citons Velupillai (2004) pour en donner une idée intuitive :

“The key word \dots is *effective*, meaning a procedure whose execution is specified in a finite series of instructions, each of which is finite in length and all the details of the execution are specified exactly, which, thus, leaves no room for magic, miracles, creative imagination and other such metaphysical entities.”

4. Les premiers décimaux d’ x sont obtenus selon la formule : $0.3984375 = \underline{0} \times 2^{-1} + \underline{1} \times 2^{-2} + \underline{1} \times 2^{-3} + \underline{0} \times 2^{-4} + \underline{0} \times 2^{-5} + \underline{1} \times 2^{-6} + \underline{1} \times 2^{-7}$.

En raison de cette correspondance stricte, s est aléatoire si et seulement si le réel lui correspondant, x est aléatoire.

La réalisation d'un test statistique consiste à déterminer de manière **effective** une série d'intervalles $I_{m,1}, I_{m,2}, \dots, I_{m,n}$ dans $[0, 1]$, où m désigne le niveau de test et n le nombre d'intervalles à éliminer au niveau m , puis vérifier si le nombre réel x se trouve dans l'un de ces intervalles $I_{m,1}, I_{m,2}, \dots, I_{m,n}$. Si tel est le cas, x ne réussit pas le test statistique ; il n'est pas aléatoire.

Ici, le terme "effectif" distingue les fonctions mathématiques abstraites de celles qui sont réalisables par un programme informatique. Plus spécifiquement, le terme "effectif" indique que les extrémités des intervalles $I_{m,1}, I_{m,2}, \dots, I_{m,n}$ forment un ensemble récursivement énumérable⁵ (désormais RE). Sans entrer dans sa définition formelle, *un ensemble est RE, si l'on peut énumérer tous ses objets avec un programme informatique.*

Par exemple, le test "la fréquence de 1 dans les premiers m termes ne dépasse pas 0,5" est un test statistique effectif qui se réalise comme suit :

Quand $m = 1$, le test examine uniquement le premier terme de s , si ce premier terme est 1, la fréquence de 1 sera 100% > 50%, s ne réussit pas le test.

Quand $m = 2$, le test étudie les premiers deux termes de la suite. Les premiers deux termes s a 4 possibilités : "00", "01", "10", "11". Selon le test, les suites commençant par "11" seront éliminées, car la fréquence de 1 dans "11" est supérieure à 0,5. Graphiquement, les suites commençant par "11" se trouvent dans le dernier quart de l'intervalle $(0, 1)$ (cf. la figure 3.1).

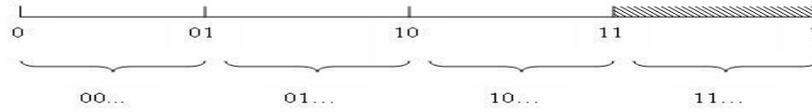
On constate dans les figures 3.1(a), 3.1(b) et 3.1(c) qu'en augmentant le niveau de test m , les parties éliminées deviennent de plus en plus fines et étalées dans l'intervalle $[0, 1]$.

Certes, avec un niveau de test modeste, on risque d'éliminer des suites aléatoires par erreur, mais quand m est suffisamment grand, ce type d'erreurs sera rectifié.

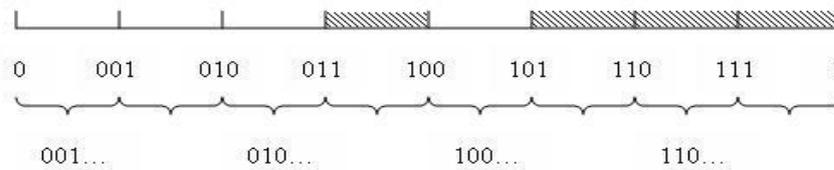
Quand $m \rightarrow \infty$, les intervalles éliminés par ces tests effectifs seront de longueur négligeable ; ils constituent, alors ce qu'on appelle un ensemble de mesure nulle dont la définition est donnée dans l'annexe K.4.

5. l'annexe K donne une définition détaillée des ensembles récursivement énumérables.

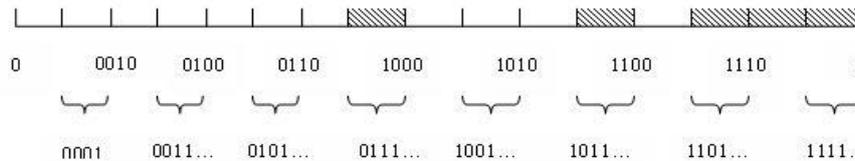
3.3. Une approche algorithmique



(a) Exemple test effectif ($m=2$), la partie ombrée est à éliminer.



(b) Exemple test effectif ($m=3$), la partie ombrée est à éliminer.



(c) Exemple test effectif ($m=4$), la partie ombrée est à éliminer.

Fig. 3.1: 3 premières étapes du test effectif “la fréquence de 1 dans les premiers m termes ne dépasse pas 0,5”.

Pour donner une idée intuitive, la mesure d’un objet indique sa fréquence, sa longueur ou son volume par rapport à un ensemble d’objets donné. Par exemple, la mesure d’un intervalle $[a, b]$ dans l’intervalle $[0, 1]$ est la longueur de $[a, b]$, donnée par $b - a$. Un ensemble d’intervalles $I_{m,1}, I_{m,2}, \dots, I_{m,n}$ est de mesure nulle si sa longueur totale tend vers 0 quand le niveau du test m augmente.

Par exemple, l’ensemble de Cantor, introduit par le mathématicien allemand Georg Cantor en 1883, est un ensemble de mesure nulle. On peut le construire de manière itérative à partir du segment $[0, 1]$ en enlevant le tiers central, puis on répète l’opération sur les deux segments restants, et ainsi de suite. On peut voir les sept premières itérations du procédé dans la figure 3.2.

On a donc :

$$\begin{aligned}
 A_0 &= [0, 1] \\
 A_1 &= [0, \frac{1}{3}] \cup [\frac{2}{3}, 1] \\
 A_2 &= [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1] \\
 &\dots
 \end{aligned}$$

L'ensemble de Cantor est l'intersection de A_n quand n tend vers l'infini.

Pour calculer sa mesure, on peut couvrir A_1 avec un ensemble des deux intervalles de mesure $\frac{2}{3}$, A_2 avec un ensemble des quatre intervalles de mesure $(\frac{2}{3})^2$. Nous pouvons donc couvrir A_n avec un ensemble des intervalles de mesure $(\frac{2}{3})^n$. Quand n tend vers ∞ , la mesure de l'ensemble des intervalles utilisés tend vers 0.



Fig. 3.2: Construction de l'ensemble Cantor

Étant donnée la correspondance entre les tests statistiques effectifs et les ensembles de mesure nulle RE, on peut définir les suites aléatoires au sens de Martin-Löf (1966) comme suit :

Définition “Suite aléatoire au sens de Martin-Löf” 3.3.1.1

Une suite infinie correspondant au nombre réel x est aléatoire au sens de Martin-Löf, si pour toute suite récursivement énumérable infinie A_i d'ensembles d'intervalles, chaque A_i de mesure majorée par 2^{-i} ($\mu(A_i) < 2^{-i}$) : x n'appartient pas à au moins un des A_i .

Remarquons qu'ici, la mesure de l'ensemble A_i diminue avec le niveau de test i . La condition ($\mu(A_i) < 2^{-i}$) impose que la mesure d' A_i tende vers 0 avec i .

Comme les suites régulières correspondent aux ensembles de mesure nulle RE, leur proportion dans l'ensemble des suites infinies est égale à 0. Par conséquent, non seulement les suites aléatoires au sens de Martin-Löf existent, mais celles-ci sont même infiniment plus nombreuses que les suites régulières.

3.3.2 *Hasard et stratégies de pari*

En pariant sur les termes futurs d'une suite aléatoire, nul ne peut gagner continuellement. Ce principe qui rejoint plus ou moins le concept du marché imbattable, a été exploité avec succès par Schnorr (1971), en appliquant la notion de "martingale" proposée par le célèbre mathématicien français Levy (1937).

Une martingale est une fonction, $f(a_1, a_2, \dots, a_n) \geq 0$ qui désigne la richesse engendrée par une stratégie donnée en pariant sur les premiers n instants d'une suite infinie composée de 0 et de 1. Ici, a_1, a_2, \dots, a_n désignent les premiers n termes de la suite binaire considérée.

Par exemple, à l'instant n , selon les termes historiques a_1, a_2, \dots, a_n , la stratégie essaie de prévoir le prochain terme a_{n+1} , et en fonction de cette prévision, elle décide s'il faut parier sur a_{n+1} , si oui, sur quel symbole (0 ou 1) et combien sera la mise M .

À l'instant $n + 1$, le terme a_{n+1} se révèle, nous envisageons trois résultats éventuels :

- La stratégie a parié sur le bon symbole, nous avons dans ce cas,

$$f(a_1, a_2, \dots, a_{n+1}) = f(a_1, a_2, \dots, a_n) + M.$$

- La stratégie a parié sur le mauvais symbole, alors,

$$f(a_1, a_2, \dots, a_{n+1}) = f(a_1, a_2, \dots, a_n) - M.$$

- La stratégie a décidé de ne pas parier,

$$f(a_1, a_2, \dots, a_{n+1}) = f(a_1, a_2, \dots, a_n).$$

Ainsi, la trajectoire de la martingale $f(a_1, a_2, \dots, a_n)$ dépend de la qualité des prévisions. Selon Levy (1937),

Définition "Battre un nombre réel" 3.3.2.1

Une suite $a_1, a_2, \dots, a_n, \dots$ (ainsi que le nombre réel lui correspondant) est **battue** par une martingale $f(\cdot)$, si

$$\limsup_{n \rightarrow \infty} f(a_1, a_2, \dots, a_n) = \infty$$

Schnorr (1971) a introduit la notion de la calculabilité dans la définition de la martingale. Selon cet auteur :

Définition “Suite aléatoire au sens de Schnorr” 3.3.2.2

Une suite infinie est aléatoire, si elle n'est battue par aucune martingale effective.

Ici, le terme “effective” signifie qu'à chaque étape n , $f(a_1, a_2, \dots, a_n)$ peut être calculé par une machine de Turing, un dispositif bien défini en informatique capable de réaliser toutes les procédures effectives (un exemple de la machine de Turing est donnée dans l'annexe K).

La définition de Schnorr est illustrée par l'exemple suivant :

Soit $a_1, a_2, \dots, a_n, \dots$, une suite de 0 et de 1 vérifiant la régularité “tous les dixièmes termes sont 0”. $a_1, a_2, \dots, a_n, \dots$ est battable par la stratégie effective : “parier sur 0 à chaque neuvième instant”.

Sans perdre de généralité, commençons la stratégie avec un capital initial de 10 euro. En ne faisant aucune opération pendant les 9 premiers termes, on obtient :

$$f(a_1) = f(a_1, a_2) = \dots = f(a_1, a_2, \dots, a_9) = 10$$

À l'instant 9, la stratégie parie en 0 sur le dixième terme avec une mise de $M = 2$ euros. Donc, à l'instant 10, on gagne 2 euros, et $f(a_1, a_2, \dots, a_9, a_{10}) = 12$. $f(\cdot)$ ne change pas entre le 11ième et 19ième terme, pour augmenter du 2 au 20ième terme, et ainsi de suite.

La limite de $f(a_1, a_2, \dots, a_n)$ tend donc vers $+\infty$, comme illustré dans la figure 3.3. Ainsi, la martingale effective $f(\cdot)$ est une stratégie de pari qui bat la suite $a_1, a_2, \dots, a_n, \dots$.

La correspondance “imprévisibilité-imbattabilité” mise en avant par la définition de Schnorr est particulièrement intéressante pour les chercheurs en finance. Le principe “imprévisibilité - imbattabilité” que nous avons défendu dans le premier chapitre semble partager la même logique fondamentale. Malheureusement, la définition des suites imbattables présentée dans cette section ne s'applique pas directement aux suites de rentabilités financières, car elle concerne uniquement les suites *infinies composées de 0 et de 1*. Pour adapter cette définition aux données financières, nous allons généraliser la définition de Schnorr dans le dernier chapitre de cette thèse en définissant les séries de cours imbattables.

3.3. Une approche algorithmique

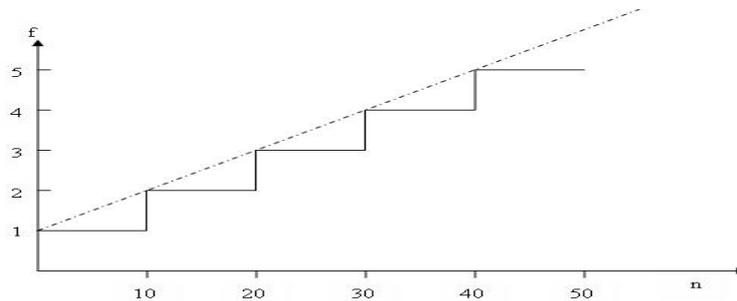


Fig. 3.3: Trajectoire de $f(\cdot)$ en considérant la stratégie “parier sur 0 à chaque neuvième instant.” L’axe Y désigne la richesse $f(\cdot)$, l’axe X désigne le temps n . Ici, on constate que $\lim_{n \rightarrow \infty} f(a_1, a_2, \dots, a_n) = \infty$

3.3.3 Hasard et programmes de compression

Considérons les trois suites *infinies* ci-dessous représentées par leurs premiers termes :

A : 01010101010101010101010101010101...

B : 01101001100101101001011001101001...

C : 11001010110000011000010111011011...

Comment les reproduire avec un programme informatique ? Pour résoudre ce problème, nous disposons des informations suivantes :

- La suite *A* est une répétition de 01.
- La suite *B* est constituée selon le principe suivant : commençons par le premier terme 0, en lui attachant son complément 1, on obtient les deux premiers termes 01. Ensuite, on y rajoute le complément de 01 pour obtenir 0110. En répétant l’opération une troisième fois, on obtient 01101001, et ainsi de suite.
- La suite *C* est le résultat enregistré d’une série de lancers d’une pièce équilibrée, ou de jeux de “pile ou face”. On note “pile” par 1 et “face” par 0.

La représentation informatique d’une suite donnée est souvent fondée sur les régularités décrites ci-dessus. Dans notre exemple,

- Pour engendrer les 2 premiers millions de termes de la suite *A* : il suffit d’imprimer un million de fois la séquence 01. Dans la plupart des langages de programmation, cela est réalisable par un petit programme composé d’une vingtaine de caractères.

Par exemple, le code écrit en langage de programmation “R” (R Development Core Team 2005) pour effectuer une telle opération se présente comme :

```
> print(makeNstr("01", 1000000))
```

- La suite B est plus difficile à engendrer dans la mesure où sa régularité inhérente est plus complexe à décrire. Cela étant, il n’est pas difficile de concevoir un programme de longueur raisonnable qui engendre les 2 premiers millions de termes de B en s’appuyant sur sa régularité structurelle. En “R”, un tel programme s’écrit comme :

```
> a="0"
> for (i in 1:(log(2000000)/log(2)))
> {
> x=chartr("0","2",a)
> x=chartr("1","0",x)
> x=chartr("2","1",x)
> a=c(a,x)
> }
> print(a)
```

- Pour engendrer les 2 premiers millions de termes de la suite C qui ne contient aucune régularité, nous sommes obligés de taper terme par terme avec une commande du type “print 110010...”. Ce programme contient au moins 2 millions de caractères.

En “R” cela donne :

```
> print("11001010110000011000010111011011...")
```

Remarquons que pour les suites A et B , nous avons réussi à réduire leur taille (2 millions) à des programmes informatiques de longueur inférieure à 100 caractères. Nous avons en quelque sorte “**comprimé**” les deux suites grâce à leur structure respective.

Le relation “régularité-compression”, illustrée dans l’exemple précédent, est en effet le principe défendu dans la complexité de Kolmogorov, le fondement théorique de la troisième définition des suites aléatoires.

Pour donner définir la complexité de Kolmogorov, nous devons introduire la notion de “machine de Turing universelle”. Une machine de Turing universelle est une machine de Turing capable de simuler toutes les machines de Turing⁶. Comme une machine de Turing universelle, U peut réaliser tous les procédures effectives, on peut donc engendrer toutes les suites finies par U .

Définition “Complexité de Kolmogorov” (Kolmogorov 1965) 3.3.3.1

Le contenu en information d’une suite finie s , ou la complexité de Kolmogorov de s , noté $K(s)$, est la taille du plus petit programme capable de produire s .

6. Pour plus de détails concernant les machines de Turing, on renvoie les lecteurs à l’annexe K

Comme nous l'avons illustré dans la comparaison des suites A , B et C au début de cette section, la complexité de Kolmogorov, $K(s)$ d'une suite donnée mesure le degré d'aléa d' s . D'une manière générale, moins s est ordonnée, plus grande sera sa complexité de Kolmogorov $K(s)$.

Remarquons ici que la complexité de Kolmogorov est définie à l'aide d'une machine de Turing universelle U . Comme il existe un nombre infini de machines de Turing universelles, l'indicateur d'aléa manquerait de fiabilité, si $K(s)$ varie radicalement selon la machine universelle utilisée. Ce problème est résolu par le théorème d'invariance, d'abord démontré par Solomonoff (1960), puis indépendamment par Kolmogorov (1965) et Chaitin (1969).

Notons $K_U(s)$ la complexité de Kolmogorov de la suite s obtenue en utilisant la machine de Turing universelle U , et $K_V(s)$ celle obtenue avec V , alors,

Théorème “d'invariance” 3.3.3.2

Il existe une constante c_{UV} , qui ne dépend que de U et V , telle que pour toute suite finie s : $|K_U(s) - K_V(s)| < c_{UV}$

Ainsi, la complexité de Kolmogorov d'une suite donnée est indépendante, à une constante additive près, de la machine de Turing universelle utilisée. Prenant comme indicateur d'aléa la complexité de Kolmogorov, Chaitin (1987) définit les suites aléatoires comme suit :

Définition “Suite aléatoire” 3.3.3.3

Une suite infinie composée de 0 et de 1, notée s , est dite aléatoire si la complexité de Kolmogorov de ses n premiers termes, notée $K^n(s)$, satisfait,

$$\exists c \forall n K^n(s) \geq n - c. \tag{3.1}$$

Où “ c ” est un paramètre fini.

Ainsi, en augmentant le nombre des termes considérés (i.e. n), le meilleur taux de compression ⁷ qu'on peut réaliser sur une suite aléatoire tend vers 0.

Selon cette définition, nous pouvons distinguer les suites aléatoires des suites ordonnées à l'aide des algorithmes de compression. Quand n est suffisamment grand, plus les premiers n termes d'une suite infinie s sont compressibles, moins s sera aléatoire.

7. Le taux de compression est un indicateur en informatique qui mesure la capacité d'un programme de compression à réduire la taille d'un fichier donné. Le taux de compression réalisé par sur une suite s est calculé par le ratio $\frac{\text{taille d's avant compression}}{\text{taille d's après compression}} - 1$.

Bien sûr, les tests de compression sont toujours conditionnés par la disponibilité des données, n ne peut pas être aussi grand que nous le souhaiterions. Par conséquent, les conclusions tirées des tests de compression sont toujours soumises à une marge d'erreur qui se réduit en augmentant la longueur des suites analysées.

Si la théorie des probabilités mesure le degré d'aléa avec l'indicateur "p-value", l'approche algorithmique peut estimer la complexité d'une suite donnée avec le taux de compression maximal qu'on peut obtenir. L'objectif du chapitre présent est donc d'introduire cette approche dans les analyses des séries financières, et de montrer à travers des exemples concrets l'intérêt d'une telle introduction.

Conclusion

Les trois définitions de suites aléatoires rejoignent notre vision du hasard exposée dans l'introduction du chapitre : le hasard défini comme les éléments incompréhensibles dans un système de préconceptions donné, est une propriété inhérente du système concerné, qui est indépendante de notre capacité de déduction.

Les suites aléatoires au sens de Martin-Löf existent, elles sont mêmes infiniment plus nombreuses que les suites régulières. Les nombres réels entre 0 et 1 éliminés par les tests effectifs, correspondant donc aux suites régulières, forment un ensemble de mesure nulle ; leur nombre est donc négligeable par rapport à l'ensemble des nombres réels dans l'intervalle $[0, 1]$ dont la mesure est égale à 1. La majorité des suites binaires infinies sont donc aléatoires au sens de Martin-Löf.

Un point essentiel concernant les suites aléatoires reste la parfaite équivalence entre les trois définitions présentées. Schnorr (1971) a montré sur le plan mathématique qu'une suite aléatoire dans le sens de Martin-Löf (1966) l'est également dans le sens de Schnorr (1971) ou de Chaitin (1987).

- Autrement dit, si une suite infinie composée de 0 et de 1 est aléatoire, alors,
- soit elle réussit tous les tests statistiques effectifs,
 - soit elle n'est battable par aucune martingale effective,
 - soit son taux de compression tend vers 0 quel que soit l'algorithme de compression utilisé.

L'équivalence entre les tests statistiques et les stratégies de pari est relativement familière aux chercheurs en finance de marché. Depuis Alexander (1961), la théorie de l'efficacité a toujours été vérifiée sous deux angles de vue différents : les tests de propriétés statistiques et les tests des stratégies d'investissement.

En revanche, la compressibilité des séries financières reste une notion originale en finance. Ce chapitre vise donc à introduire une nouvelle approche empirique dans les études du hasard financier, en explorant l'équivalence entre **l'incompressibilité** et **l'imbattabilité**. Plus précisément, la complexité de Kolmogorov des séries financières sera estimée à l'aide des algorithmes de compression.

Or, comme nous l'avons évoqué précédemment, les résultats théoriques obtenus en informatique fondamentale sont valables uniquement sur les suites infinies composées de 0 et de 1. L'usage des nombres réels dans la représentation des rentabilités financières rend impossible une application directe des outils informatiques. C'est pourquoi nous proposerons une méthode générale permettant de transformer les rentabilités financières en séries de nombres entiers qui sont, elles, informatiquement manipulables.

3.4 Les algorithmes de compression : des détecteurs de régularités

La complexité de Kolmogorov d'une suite donnée, définie comme le plus petit programme qui l'engendre, est un indicateur fiable du hasard algorithmique. Elle est par définition la limite théorique de tous les algorithmes de compression.

Cependant, la complexité de Kolmogorov d'une suite donnée n'est pas directement calculable. En pratique, les chercheurs en informatique utilisent souvent les algorithmes de compression pour estimer la complexité des données. La logique de cette démarche peut être résumée comme suit : puisque la complexité de Kolmogorov d'une suite "s" est la limite théorique de tous les algorithmes de compression, nous pouvons estimer cette complexité avec le meilleur taux de compression réalisé en appliquant n différents algorithmes de compression sur "s". Et cette estimation est d'autant plus fiable que la valeur de n augmente.

Inspirée par cette pratique, nous allons proposer une approche empirique pour estimer la complexité de Kolmogorov des séries financières.

Pour cela, le choix de la méthode de compression est essentiel au résultat final de l'opération, car tous les algorithmes de compression ne cherchent pas les mêmes régularités : les données incompressibles par un algorithme spécifique peuvent s'avérer compressibles par un autre.

Il est donc important de connaître le mode de fonctionnement des principales méthodes de compression afin de mieux interpréter les résultats obtenus avec les séries financières.

À cet effet, deux grandes catégories d’algorithmes de compression sont présentées dans cette section : le codage d’entropie qui exploite la fréquence d’apparition des symboles et le codage par contexte qui s’inscrit davantage dans une logique de prédiction. C’est parmi ces deux catégories d’algorithmes, reconnues comme les plus performantes pour des compressions sans perte, que nous allons choisir “nos détecteurs de régularités financières”.

Le choix de ces deux catégories d’algorithmes est aussi fondé sur leur réversibilité. Autrement dit, en appliquant ces algorithmes sur une suite s , on peut retrouver à partir du résultat obtenu, la série d’origine s sans aucune perte d’information. Ces algorithmes de compression sans perte sont souvent utilisés dans les traitements de textes, car il est essentiel pour la compréhension d’un texte donné qu’aucun mot ne soit omis pendant la procédure de compression. En revanche, cela n’est pas le cas pour les traitements d’images ou de sons où on peut parfois consacrer la qualité d’une image (ou d’un son) pour obtenir un meilleur résultat de compression.

Pour analyser des séries financières, nous estimons qu’il est essentiel de conserver toutes les informations contenues dans les données historiques, c’est pourquoi nous avons choisi d’utiliser les méthodes de compression sans perte.

3.4.1 Codages d’entropie

La compression de données est l’opération informatique qui consiste à réduire la taille de l’information pour le stockage ainsi que son transport. Inspirées par l’entropie de Shannon présentée ci-dessous, certaines méthodes de compression cherchent à coder les différents symboles en fonction de leur fréquence d’apparition. Comme le nom l’indique, ce type d’algorithmes a pour objectif de rapprocher au maximum la longueur d’un texte donné à son entropie de Shannon. Leur principe de base est simple : les mots les plus fréquents sont codés de manière plus concise.

Par exemple, dans ce travail de thèse, on pourrait créer un dictionnaire qui au mot “aléa” associe le code “00”, au mot “finance” le code “01”, aux mots “efficience des marchés financiers” le code “10”, et ainsi de suite. On en réduirait alors la taille sans problème.

Dans cette section, nous allons d’abord rappeler avec un exemple concret la définition de l’entropie de Shannon, et puis présenter les trois catégories de codages d’entropie les plus souvent utilisées dans les compressions sans perte, à savoir, le codage dit de Huffman, le codage arithmétique et le codage par dictionnaire.

Entropie de Shannon

Selon Shannon (1948), l'entropie d'une variable aléatoire discrète, X comportant n symboles différents est définie par l'équation (3.2).

$$H(X) = E(-\log_2(Pr(X = x_i))) = \sum_{i=1}^n Pr(X = x_i) \log_2\left(\frac{1}{Pr(X = x_i)}\right) \quad (3.2)$$

Où $H(X)$ désigne l'entropie de Shannon de X , $E(.)$ signifie l'espérance mathématique, et $Pr(X = x_i)$ la probabilité d'observer $X = x_i$.

Par exemple, l'entropie de la variable X définie par l'équation (3.3)

$$X = \begin{cases} x_1 = A & Pr(X = x_1) = 60\% \\ x_2 = B & Pr(X = x_2) = 40\% \end{cases} \quad (3.3)$$

est calculée par :

$$H(X) = 60\% \log_2 \frac{1}{60\%} + 40\% \log_2 \frac{1}{40\%} = 0,97. \quad (3.4)$$

L'entropie de Shannon illustre la quantité d'information contenue **en moyenne** dans chaque symbole du message. Dans l'exemple précédent, la variable X a une entropie de $H(X) = 0,97$, cela signifie que chaque symbole de X contient en moyenne 0,97 bits⁸ d'information.

Remarquons que pour transmettre une quantité d'informations donnée, moins on utilise de symboles différents, plus le texte final paraîtra long. Par exemple, en traduisant un texte français en chinois, on obtient souvent un texte considérablement plus court. Il existe en français 26 lettres au total. En supposant que chaque lettre ait la même chance d'apparaître, le contenu en information de chaque alphabet français est de $\log_2 26 = 4.7$ bits.

En revanche, il y a 2500 caractères chinois communs, en supposant que chaque caractère ait la même chance d'apparaître, le contenu en information de chaque caractère est de $\log_2 2500 = 11.3$ bits.

Par conséquent, pour transmettre la même quantité d'information, le nombre de caractères chinois nécessaire est moins important que celui des lettres françaises. Un texte en chinois est donc moins long que son équivalent en français.

8. Le "bit" est une unité de mesure en informatique qui désigne la quantité d'information contenue dans un chiffre binaire. Un bit reçoit donc soit "0" soit "1".

Bien sûr, tous les alphabets français (ou caractères chinois) n'ont pas la même fréquence d'apparition. De ce point de vue, notre calcul est une approximation de la réalité. Néanmoins, cet exemple illustre bien la relation entre le nombre de symboles utilisés et la longueur d'un texte donné.

Puisque l'entropie de Shannon mesure la quantité d'information contenu **en moyenne** dans chaque symbole utilisé, on peut calculer la longueur minimale d'un texte s à partir de son l'entropie de Shannon, $H(x)$. En supposant que s soit de longueur n , la longueur minimale de s est de " $H(x) \times n$ ". Cette longueur minimale est parfois considérée comme la limite des compressions sans perte.

Cependant, il est nécessaire de préciser que le contenu en information calculé par l'entropie de Shannon ne coïncide pas toujours avec la complexité de Kolmogorov. Coder un texte selon la fréquence d'apparition des symboles est une méthode de compression, mais elle n'est pas la seule. Des régularités non statistiques, qui n'influencent pas toujours la fréquence d'apparition des symboles, peuvent parfois donner lieu aux meilleurs résultats de compression ⁹.

Par exemple,

- Soit $A = 01010101 \dots$ une suite engendrée en répétant 1 million de fois la séquence 01.
- Soit $B = 01110110 \dots$ une suite obtenue en lançant 2 millions de fois une pièce équilibrée.

La limite de compression prévue par l'entropie de Shannon pour ces deux suites est : $H(X) \times n = (\frac{1}{2} \log_2(2) + \frac{1}{2} \log_2(2)) \times 2000000 = 2000000$. Or, la suite A est moins complexe que B , et la compression maximale qu'on peut réaliser sur A est largement supérieure à son entropie de Shannon.

Ainsi, l'entropie de Shannon est la limite de compression pour les algorithmes qui exploitent la fréquence des symboles. En revanche, la complexité de Kolmogorov qui est la théorie la plus générale de la régularité, est la limite ultime de toutes les méthodes de compression effectives.

Il semblerait que cette nuance est ignorée par certains travaux en finance (ex. Mansilia (2004)) qui utilisent l'entropie de Shannon pour estimer la complexité de Kolmogorov des séries financières.

Après avoir éclairci ces concepts de base, nous présentons maintenant les prin-

9. Certains algorithmes ont déjà emporté le défi de l'entropie de Shannon. "Multipass holoDynammic compression" de Sandy Shaw, "bit compression" de Robert Langley et "autosophy information theory" de Klaus Holtz semblent pouvoir produire de meilleurs résultats que l'entropie de Shannon. Il existe une liste des algorithmes "éminents" sur le site <http://www.endlesscompression.com/>

cipaux algorithmes de compression visant à rapprocher la longueur d'un texte à son entropie de Shannon.

I. Codage de Huffman

Selon l'article de Huffman (1952), un codage sans perte qui minimise la longueur d'un texte donné doit remplir 5 conditions :

- Deux mots différents ne bénéficient jamais d'un code identique.
- Les digits marquant le début et la fin de chaque mot ne sont pas indispensables.
- Plus souvent un mot apparaît dans un texte, moins long serait son code correspondant.
- Au moins deux des plus longs codes doivent être différenciés uniquement par leurs derniers digits. Sinon, on peut supprimer les derniers digits des plus longs codes pour les raccourcir.
- Soit $L(n)$, le code le plus long dans le système du codage. Toutes les séquences plus courtes que $L(n)$ doivent être utilisées dans le même système, car si l'une des séquences plus courtes ne fait pas encore partie du système, alors, nous pouvons l'utiliser pour coder le mot initialement représenté par $L(n)$.

En respectant toutes ces conditions, l'auteur a proposé un codage efficace que nous pouvons illustrer avec un simple exemple :

Supposons qu'un texte donné soit composé de l'alphabet {efficience, compression, hasard, Kolmogorov}, les fréquences d'apparition des symboles sont respectivement {40%, 35%, 20%, 5%}. On crée un arbre avec les quatre symboles initiaux (efficience, compression, hasard, Kolmogorov), on prend les deux symboles les moins fréquents (hasard, Kolmogorov) pour former un autre mot dont la fréquence d'apparition est la somme des deux symboles fusionnés. Répétons l'opération jusqu'au moment où il n'existe qu'une seule branche dans l'arbre du codage, et on code l'arbre de droite à gauche avec les séquences de 0 et de 1 (cf. figure 3.4).

Les codes de Huffman utilisés pour compresser le texte initial sont présentés dans le tableau 3.1 :

II. Codage arithmétique

Une autre méthode pour approcher l'entropie de Shannon et la longueur d'un texte donné est dite du "codage arithmétique". Contrairement au codage de Huff-

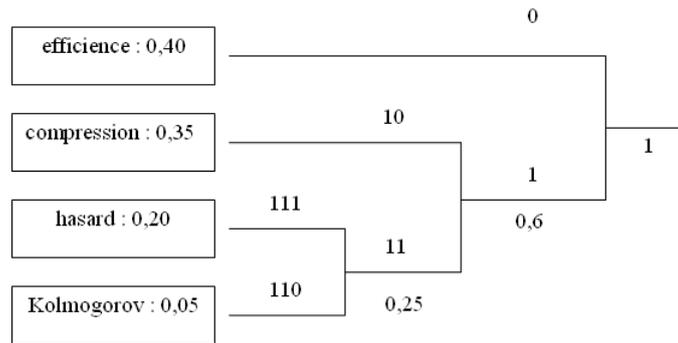


Fig. 3.4: Arbre de Huffman

alphabet	code
efficience	0
compression	10
hasard	111
Kolmogorov	110

Tab. 3.1: Code de Huffman correspondant à chaque mot du texte à compresser

man qui construit un arbre dès le départ, la méthode arithmétique code des textes mot par mot. Cela nécessite donc un modèle mathématique qui calcule, à chaque étape, la probabilité d'apparition de **chaque** symbole. Pour les textes relativement longs, la probabilité d'apparition d'un même symbole peut évoluer dans le temps.

Illustrons maintenant le principe du codage arithmétique avec un exemple concret :

Soit “NEUTRALPOSITIVENEGATIVEEND-OF-DATA” le début d'un message défini sur l'alphabet $\{\text{NEUTRAL}, \text{POSITIVE}, \text{NEGATIVE}, \text{END-OF-DATA}\}$. Supposons que sur l'ensemble du texte, la probabilité d'apparition de chaque mot de l'alphabet est respectivement $\{\frac{3}{5}, \frac{1}{5}, \frac{1}{10}, \frac{1}{10}\}$ ¹⁰.

Le codage arithmétique du message est illustré dans la figure 3.5.

1. À l'étape I, l'intervalle initial $[0, 1[$ est divisé en 4 parties selon les probabilités d'apparition de chaque mot :
 - l'intervalle du mot “NEUTRAL” est $[0, \frac{3}{5}[$,
 - celui de “POSITIVE” est $[\frac{3}{5}, \frac{3}{5} \times \frac{3}{5} + \frac{1}{5}[= [\frac{3}{5}, \frac{4}{5}[$,

¹⁰ Comme nous l'avons évoqué, les probabilités d'apparition peuvent varier dans le temps. Ici, pour illustrer le principe du codage arithmétique, nous avons pris un exemple simple où les probabilités d'apparition restent stables dans le temps.

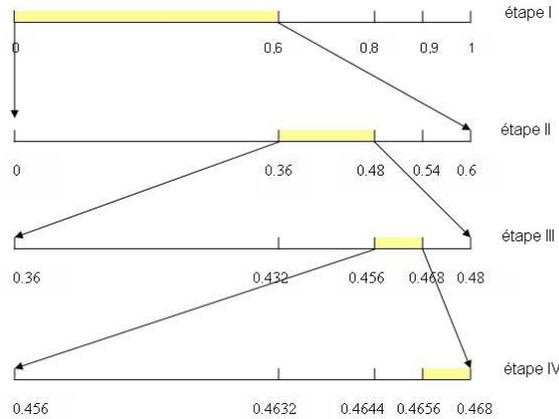


Fig. 3.5: Codage arithmétique du message “NEUTRALPOSITIVENEGATIVEEND-OF-DATA”

- celui de “NEGATIVE” est $[\frac{4}{5}, \frac{4}{5} + \frac{1}{10}[= [\frac{4}{5}, \frac{9}{10}[$,
- en fin, celui de “END-OF-DATA” est $[\frac{9}{10}, 1[$.

Comme le premier mot du message est “NEUTRAL”, le code du texte se trouvera dans l’intervalle $[0, \frac{3}{5}[$.

2. À l’étape II, l’intervalle $[0, \frac{3}{5}[$ est divisé en quatre parties selon les probabilités d’apparition, ce qui donne $[0, \frac{3}{5} \times \frac{3}{5}[$ pour “NEUTRAL”, $[\frac{3}{5} \times \frac{3}{5}, \frac{3}{5} \times (\frac{3}{5} + \frac{1}{5})[$ pour “POSITIVE”, $[\frac{3}{5} \times (\frac{3}{5} + \frac{1}{5}), \frac{3}{5} \times (\frac{3}{5} + \frac{1}{5} + \frac{1}{10})[$ pour “NEGATIVE”, et $[\frac{3}{5} \times (\frac{3}{5} + \frac{1}{5} + \frac{1}{10}), \frac{3}{5}[$ pour “END-OF-DATA”. Comme le deuxième mot du texte est “POSITIVE”, le code final se trouvera dans l’intervalle $[\frac{3}{5} \times \frac{3}{5}, \frac{3}{5} \times (\frac{3}{5} + \frac{1}{5})[= [0.36, 0.48[$ ¹¹.
3. À l’étape III, l’intervalle $[0.36, 0.48[$ sera divisé en quatre parties : $[0.36, 0.432[$, $[0.432, 0.456[$, $[0.456, 0.468[$ et $[0.468, 0.48[$. En observant le troisième mot “NEGATIVE”, nous concluons que le code final est dans l’intervalle $[0.456, 0.468[$.
4. Répétons l’opération encore une fois, l’intervalle $[0.456, 0.468[$ sera divisé en 4 parties : $[0.456, 0.4632[$, $[0.4632, 0.4644[$, $[0.4644, 0.4656[$ et $[0.4656, 0.468[$. En observant le dernier mot “END-OF-DATA”, le code final sera $[0.4656, 0.468[$. Remarquons qu’ici, tous les nombres réels dans l’intervalle $[0.4656, 0.468[$ représentent le texte “NEUTRALPOSITIVENEGATIVEEND-OF-DATA”. En augmentant la longueur du message, l’intervalle retenu sera de plus en plus réduit.

La question maintenant est de savoir comment retrouver le message initial, à partir du code $[0.4656, 0.468[$. Cette procédure nommée la “décompression” s’opère comme suit :

11. Ici, pour simplifier la notation des intervalles, nous adoptons la représentation anglophone du décimal, “.”.

1. Divisons l'intervalle $[0, 1[$ en quatre selon les probabilités d'apparition : $[0, 0.6[$, $[0.6, 0.8[$, $[0.8, 0.9[$, $[0.9, 1[$. Comme le code $[0.4656, 0.468[$ se situe dans le premier intervalle $[0, 0.6[$, on obtient le premier mot du texte, "NEUTRAL".
2. Divisons ensuite l'intervalle $[0, 0.6[$ en quatre parties : $[0, 0.36[$, $[0.36, 0.48[$, $[0.48, 0.54[$, $[0.54, 0.6[$.
Comme $[0.4656, 0.468[$ se trouve dans l'intervalle $[0.36, 0.48[$, le deuxième mot du message est "POSITIVE".
3. En répétant cette opération 2 fois de plus, on retrouve le message d'origine.

Deux points sont à préciser concernant le codage arithmétique :

- Le codage binaire est préféré, par rapport au code décimal utilisé dans l'exemple précédent, car un digit décimal peut contenir plus d'informations qu'un digit binaire. Dans notre exemple, le texte "NEUTRALPOSITIVE NEGATIVEEND-OF-DATA" peut être codé par les nombres décimaux 0.466, 0.467 ou 0.468¹² qui contiennent chacun 3 digits décimaux. D'une façon générale, trois digits décimaux peuvent contenir $H = 3\log_2(10) = 9.97$ bits d'information. En revanche, le même message peut aussi être codé par la séquence binaire 0.01110101 (0.4570313 en décimal) qui ne coûte que 8 bits.
- La redondance du code arithmétique¹³ est expliquée par la longueur du texte à coder. En effet, le début du message "NEUTRALPOSITIVENEGATIVE END-OF-DATA" est trop court pour que les probabilités d'apparition soient un meilleur choix que les fréquences observées dans le texte. En utilisant les fréquences observées dans le calcul de l'entropie (c'est-à-dire que $\frac{1}{4}$ pour chaque mot), l'entropie du message devient $4 \times \log_2 4 = 8$, qui est égale à la longueur du code arithmétique.

III. Codage par dictionnaire

La troisième catégorie des codages d'entropie est la méthode par dictionnaire. Les algorithmes LZ (LZ77, LZ78, LZW...) sont les principales applications de cette méthode. Comme l'indique son nom, la méthode par dictionnaire vise à optimiser le codage des séquences répétitives en construisant un dictionnaire adapté.

- Par exemple, la compression du message "abccbatcbo" se déroule comme suit :
- D'abord, séparer le message en différentes parties pour que chacune d'elles soit la plus petite séquence qui n'est identique avec aucune des parties

12. Ce sont les valeurs dans l'intervalle $[0.4656, 0.468[$.

13. La redondance d'un codage est définie par l'écart entre l'entropie du message d'origine ($4 \times (0.6\log_2 \frac{1}{0.6} + 0.2\log_2 \frac{1}{0.2} + 2 \times 0.1\log_2 \frac{1}{0.1}) = 6, 28$) et la longueur de son code arithmétique (8 bits).

précédentes. Le tableau 3.2 illustre la division effectuée pour le message "abccbatbco".

a	b	c	cb	at	cbo
---	---	---	----	----	-----

Tab. 3.2: "Parties" construites à partir du message "abccbatbco". La quatrième case contient la séquence *cb*, non *c*, car *c* serait identique avec la troisième case. C'est pour cette même raison que la dernière case contient *cbo*, et non *c* ou *cb*.

- Ensuite, numéroter chaque partie avec les nombres naturels. 0 correspond au symbole $\#$ qui représente le vide. Le tableau ainsi obtenu (i.e. tableau 3.3) est le dictionnaire à utiliser pour coder le message "abccbatbco".

0	1	2	3	4	5	6
$\#$	a	b	c	cb	at	cbo

Tab. 3.3: Numéroter chaque partie avec les nombres naturels.

- Puis, coder les symboles rencontrés aux étapes précédentes avec les numéros entiers qui leur correspondent (cf. 3.4).

1	2	3	4	5	6
0a	0b	0c	3b	1t	4o

Tab. 3.4: Codage des symboles rencontrés aux étapes précédentes avec les numéros entiers qui leur correspondent

- Coder le numéro entier dans la *n*ième case en $\lfloor \log_2(n) \rfloor$ bits. Ici, " $\lfloor \log_2(n) \rfloor$ " désigne le plus grand entier qui ne dépasse pas $\log_2(n)$.
- Coder les lettres restant par leur code `ascii`¹⁴. Nous constatons dans le tableau 3.5 que le message "abccbatbco" est représenté par le code "a 0b 00c 11b 001t 100o" qui mesure $(1 + 2 + 2 + 3 + 3 + 8 \times 6) = 59$ bits. Alors avant la compression, le message comptait $8 \times 10 = 80$ bits.

Pour décompresser le code "a0b00c11b001t100o", il suffit de connaître l'arbre de décompression montré dans la figure 3.6 :

- Nous pouvons constater dans l'arbre 3.6 que la première case non vide du message contient la lettre "a". La deuxième et troisième cases contiennent respectivement "b" et "c".
- La quatrième case représente le mot "cb", la cinquième "at", et la sixième désigne "cbo".

14. Le système du code `ascii` (American Standard Code for Information Interchange), est la norme de codage de caractères en informatique la plus connue, la plus ancienne et la plus largement compatible. ASCII contient les caractères nécessaires pour écrire en anglais. Au départ, `ascii` contient 128 caractères numérotés de 0 à 127, chaque caractère est donc codé sur 7 bits. Pour coder les numéros entre 128-255, l'ISO 8859-1, aussi appelée Latin-1, étend l'`ascii` avec les caractères accentués utiles aux langues originaires d'Europe occidentale comme le français ou l'allemand.

3.4. Les algorithmes de compression : des détecteurs de régularités

1	2	3	4	5	6
$\underbrace{0}_{0bit} a \rightarrow a$	$\underbrace{0}_{1bit} b \rightarrow 0b$	$\underbrace{0}_{2bits} c \rightarrow 00c$	$\underbrace{3}_{2bits} b \rightarrow 11b$	$\underbrace{1}_{3bits} t \rightarrow 001t$	$\underbrace{4}_{3bits} o \rightarrow 100o$

Tab. 3.5: Le contenu de la n ème case est codé en “ $\lceil \log_2(n) \rceil$ ” bits.

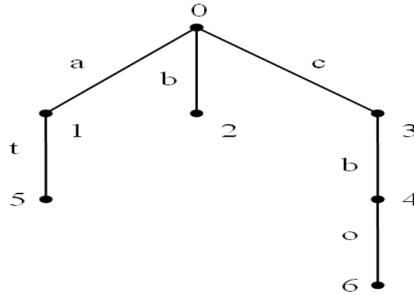


Fig. 3.6: Arbre pour décompresser un codage par dictionnaire

- En réunissant les 6 cases, on retrouve le message d’origine “abccbatcbo”. Le principe que nous avons illustré ici est adopté par tous les algorithmes de famille LZ. La seule différence qui existe entre LZ77, LZ78 et LZW se trouve dans leur manière de gérer le dictionnaire (i.e. le tableau 3.3 dans notre exemple).
- LZ78 utilise un dictionnaire de taille fixe, souvent de 16, 32 ou 64 bits, une gestion par piles est pratiquée pour organiser les entrées et les sorties des mots du dictionnaire.
- LZ77 gère son dictionnaire de façon plus dynamique, au lieu de compresser toutes les informations en une seule fois, l’algorithme utilise une fenêtre glissante et construit son dictionnaire uniquement en fonction des informations se trouvant dans la fenêtre. Cette méthode permet de mieux exploiter les régularités locales.
- LZW est une amélioration de LZ78 qui dispose d’un dictionnaire de taille évolutive. Grâce à cette flexibilité, les codes des séquences répétitives évoluent selon les propriétés locales afin de réduire au maximum la taille du fichier compressé.

Dans cette section, trois types de codage d’entropie ont été présentés : le codage de Huffman, le codage arithmétique et le codage par dictionnaire. Quelle que soit la technique utilisée, les codages d’entropie ont pour vocation d’approcher la taille d’un fichier à son entropie de Shannon. Par définition, l’entropie de Shannon est la limite théorique de ces méthodes de compression.

Or, comme nous l’avons évoqué précédemment, les codages d’entropie exploitent uniquement les régularités statistiques, i.e., les régularités indépendantes de la fréquence d’apparition des symboles. Il sera donc tout à fait possible qu’un algorithme fondé sur des régularités non-statistiques puisse engendrer de meilleurs résultats de compression.

Dans la pratique, les codages d'entropie sont souvent combinés avec d'autres techniques de compression destinées à exploiter les régularités structurelles des textes. Ces dernières techniques seront présentées dans la prochaine section.

3.4.2 Codages par contexte : algorithmes de compression dotés d'une intelligence artificielle.

Comme nous l'avons présenté précédemment, un codage arithmétique efficace implique une bonne estimation des probabilités d'apparition. Plus le modèle d'estimation est proche de la réalité, meilleur sera le résultat de compression.

Avant les années 90, les algorithmes de compression estimaient les probabilités d'apparition selon la fréquence de chaque symbole dans l'ensemble du texte. Or, en réalité, les données sont rarement homogènes, elles représentent presque toujours des particularités locales qui peuvent être exploitées par les méthodes de compression. L'usage des probabilités évolutives améliore donc considérablement les résultats de compression.

Les modèles de prédiction sont donc "rajoutés" au codage arithmétique pour calculer les probabilités d'apparition selon le "contexte local". Cette nouvelle catégorie d'algorithmes, nommée "codage par contexte", offre souvent des performances exceptionnelles en matière des compressions sans perte.

Malgré sa première apparition académique en 1984 (Cleary et Witten 1984), la mise en pratique du "codage par contexte" n'a commencé qu'à partir des années 90. À notre connaissance, il existe aujourd'hui trois principales familles d'algorithmes de cette nature : le "Partial Prediction Matching" (désormais PPM), le "Dynamic Markov Compression" (désormais DMC) et le "Context Mixing" (désormais CM). Ces algorithmes "intelligents" qui prennent leurs décisions selon des contextes locaux, sont nos principaux outils de compression dans les analyses des séries financières.

I. Partial Prediction Matching (PPM)

La méthode PPM est avant tout une amélioration du codage arithmétique. Elle se distingue d'un codage arithmétique basique dans la mesure où le PPM estime la probabilité d'apparition de chaque symbole en fonction des contextes locaux.

Plus précisément, PPM divise le codage de chaque symbole en deux grandes étapes : l'estimation des probabilités d'apparition et le codage arithmétique selon les probabilités estimées.

Par exemple, le codage du symbole (C) dans le message ci-dessous se déroule comme suit :

“ $ABCABDCAB(C)$ ”

Ici, les parenthèses autour de (C) signifient que “ C ” est le symbole à coder.

- Dans un premier temps, l'algorithme décide la taille des plus longs contextes à considérer, nommée l'ordre du PPM. Un PPM d'ordre 5, noté PPM(5), signifie que les plus longs contextes à considérer contiennent 5 symboles. D'ailleurs, “5” est souvent considéré comme l'ordre optimal d'un PPM¹⁵ (Cleary et Witten 1984, p3).
- Supposons que l'ordre optimal pour compresser le message “ $ABCABDCAB(C)$ ” soit 2, alors pour coder le symbole “ C ” entre parenthèses, le PPM(2) cherche, parmi les symboles historiques “ $ABCABDCAB$ ”, toutes les séquences de taille inférieure ou égale à 2 et les mémoriser en tant que contextes. Puis, pour chaque contexte trouvé, PPM(2) compte les différents symboles qui le suivent et calcule la fréquence d'apparition de chacun d'entre eux. Dans notre exemple, après le contexte “ AB ” on constate 1 “ D ” et 1 “ C ”.

Le PPM suppose également que pour chaque prédiction sous un contexte donné il y a un risque d'échec. Un contexte connaît un échec si aucune prédiction sous ce contexte ne correspond au vrai symbole à coder. Dans notre exemple, le contexte “ AB ” n'est pas en échec, car le symbole à coder “ C ” est l'une des prédictions sous le contexte “ AB ”. En revanche, si le symbole à coder était “ A ”, le contexte “ AB ” connaîtrait un échec, l'algorithme regarderait le contexte d'une taille inférieure (le contexte “ B ” dans notre exemple). Le passage “ $AB \rightarrow B$ ” est nommé “*escaping*” en anglais.

En supposant que chaque prédiction a autant de chance de se réaliser que de connaître un échec, les probabilités d'apparition pour coder le symbole “(C)” sont : $\frac{1}{4}$ pour “ C ”, $\frac{1}{4}$ pour “ D ”, et $\frac{1}{2}$ pour l'*escaping* (cf. tableau 3.6 page 127).

- Après l'obtention des probabilités d'apparition, le symbole “(C)” sera codé selon la méthode arithmétique présentée dans la section 3.4.1 page 118.

Remarquons que dans notre exemple, la lettre “ C ” est bien prise en compte par les contextes de taille 2. Or, si le symbole à coder était “ A ”, alors, les contextes de taille 2 connaîtraient un échec, nous devrions passer aux

¹⁵. Certains PPM récents estiment l'ordre optimal en fonction des données à compresser. Ici, nous ne rentrons pas dans les détails d'une telle estimation, pour bien nous concentrer sur le principe des compressions par contexte.

3.4. Les algorithmes de compression : des détecteurs de régularités

contextes de taille 1 grâce à la possibilité d'*escaping*. Nous pouvons remarquer dans le tableau 3.6 que même sous le contexte "B", le symbole "A" n'est pas pris en compte. Un deuxième *escaping* nous amène aux contextes de taille 0, les probabilités d'apparition pour coder "A" seront donc calculées selon la figure 3.7.

$$\left\{ \begin{array}{l} AB \rightarrow C \\ AB \rightarrow D \\ \\ AB \rightarrow Esc \end{array} \right. \frac{1}{2} \times \left\{ \begin{array}{l} B \rightarrow C \\ B \rightarrow D \\ \\ B \rightarrow Esc \end{array} \right. \frac{1}{4} \times \left\{ \begin{array}{l} \rightarrow A \\ \rightarrow B \\ \rightarrow C \\ \rightarrow D \\ \rightarrow Esc \end{array} \right. \begin{array}{l} \frac{1}{4} \\ \frac{1}{4} \\ \frac{3}{13} \\ \frac{3}{13} \\ \frac{2}{13} \\ \frac{1}{13} \\ \frac{1}{4} \\ \frac{1}{13} \end{array}$$

Fig. 3.7: Calcul des probabilités d'apparition, si le message à coder était "ABCABD-CAB(A)".

Les contextes de taille -1 seraient utilisés si le symbole à coder n'est jamais apparu dans les symboles historiques (ex. si le message à coder était "ABCABD-CAB(H)"). Dans ce cas de figure, nous imposons la même probabilité d'apparition pour chaque lettre observée dans le message entier. Les contextes de taille -1, représentés par " $\rightarrow s \ 1 \ \frac{1}{|s|}$ " dans la dernière colonne du tableau 3.6 signifie donc :

$$\begin{array}{l} \rightarrow A \ 1 \ \frac{1}{5} \\ \rightarrow B \ 1 \ \frac{1}{5} \\ \rightarrow C \ 1 \ \frac{1}{5} \\ \rightarrow D \ 1 \ \frac{1}{5} \\ \rightarrow H \ 1 \ \frac{1}{5} \end{array}$$

3.4. Les algorithmes de compression : des détecteurs de régularités

Taille 2		Taille 1		Taille 0		Taille -1	
AB→C	2 $\frac{1}{5}$	A→B	3 $\frac{3}{4}$	→A	3 $\frac{3}{13}$	→s	1 $\frac{1}{ s }$
AB→D	1 $\frac{1}{5}$	A→Esc	1 $\frac{1}{4}$	→B	3 $\frac{3}{13}$		
AB→Esc	2 $\frac{3}{5}$			→C	2 $\frac{2}{13}$		
				→D	1 $\frac{1}{13}$		
BC→A	1 $\frac{1}{2}$	B→C	1 $\frac{1}{4}$	→Esc	4 $\frac{4}{13}$		
BC→Esc	1 $\frac{1}{2}$	B→D	1 $\frac{1}{4}$				
		B→Esc	2 $\frac{1}{2}$				
BD→C	1 $\frac{1}{2}$	C→A	2 $\frac{2}{3}$				
BD→Esc	1 $\frac{1}{2}$	C→Esc	1 $\frac{1}{3}$				
CA→B	2 $\frac{2}{3}$	D→C	1 $\frac{1}{2}$				
CA→Esc	1 $\frac{1}{3}$	D→Esc	1 $\frac{1}{2}$				
DC→A	1 $\frac{1}{2}$						
DC→Esc	1 $\frac{1}{2}$						

Tab. 3.6: Les modèles de prévision créés par un algorithme PPM(2) pour le message “ABCABDCAB(C)”. Ici, s représente toutes les lettres apparues dans le message à coder, et $|s|$ désigne le nombre total de ces caractères.

II. Dynamic Markov Compression (DMC)

Fondé sur le même fonctionnement que le PPM, la *dynamic markov compression* (désormais DMC) est un autre codage par contexte. La différence entre le PPM et la DMC tient au fait que le PPM code à chaque étape un symbole occupant 8 bits, alors que la DMC traite les informations bit par bit.

Dans l'exemple précédent, le PPM lit à chaque fois un caractère (“A”, “B”, “C” ou “D”), chacun codé sur 8 bits, avant la compression. En revanche, pour traiter le même message, la DMC lit 1 seul bit à la fois, soit 1 soit 0. En conséquence, pour effectuer un codage arithmétique, il suffit d'estimer les probabilités d'apparition des 1 et des 0 à chaque bit, comme une chaîne de Markov à deux états alternatifs.

III. Context Mixing (CM)

La méthode “*context mixing*” (désormais CM) traite aussi les informations bit par bit. Sa particularité se trouve dans la combinaison des contextes dans le modèle de prédiction, d'où vient la dénomination “*context mixing*”. La famille des

compresseurs “PAQ”¹⁶, est le principal représentant des algorithmes de CM.

Comme il sera montré dans les sections prochaines, c’est cette catégorie d’algorithmes qui réalise les meilleurs taux de compression sur les séries financières analysées. Nous allons donc étudier en détail leur mode de fonctionnement pour mieux interpréter nos résultats empiriques.

Pour coder le “(1)” dans le message “1110110000101(1)...”, le CM estime les probabilités d’apparition de chaque symbole comme suit :

- En considérant comme un contexte les 8 bits devant “(1)” (soulignés dans le message initial), le CM obtient $n_{1,8} = 3$ et $n_{0,8} = 5$. Ici, $n_{1,8} = 3$ signifie “1” apparaît 3 fois dans le contexte de 8 bits et $n_{0,8} = 5$ indique que “0” apparaît 5 fois.
- En considérant les 7 bits devant “(1)” comme un contexte, on a $n_{1,7} = 2, n_{0,7} = 5$.
- En considérant les 6 bits devant “(1)” comme un contexte, on obtient $n_{1,6} = 2, n_{0,6} = 4$.
- ...

Les probabilités de “1” et de “0” sont estimées selon les équations suivantes :

$$s_1 = \sum_{i=1}^8 w_i n_{1,i} \quad (3.5)$$

$$s_0 = \sum_{i=1}^8 w_i n_{0,i} \quad (3.6)$$

$$P_1 = \frac{s_1}{s_1 + s_0} \quad (3.7)$$

$$P_0 = \frac{s_0}{s_1 + s_0} \quad (3.8)$$

Où P_1 est la probabilité d’observer “1” à la prochaine étape, et w_i désigne la pondération associée au contexte “i”. Généralement, w_i est d’autant plus grand que la longueur du contexte notée par “i” est elle même grande.

À partir de la 7ième génération de PAQ (PAQ7 délivré en 2005 par Matt Mahoney), les pondérations w_i sont fixées par un processus d’apprentissage (“*neutral*”

16. Disponibles sur le site <http://www.compressionmax.com/paq.html>

work”) défini comme suit :

$$P_{i,1} = \frac{n_{1,i}}{n_{1,i} + n_{0,i}} \quad (3.9)$$

$$P_{i,0} = \frac{n_{0,i}}{n_{1,i} + n_{0,i}} \quad (3.10)$$

$$x_i = \text{stretch}(P_{i,1}) \quad (3.11)$$

$$P(1) = \text{squash}\left(\sum_{i=1}^8 w_i x_i\right) \quad (3.12)$$

$$(3.13)$$

Avec :

$$\text{stretch}(x) = \ln\left(\frac{x}{1-x}\right) \quad (3.14)$$

$$\text{squash}(x) = \frac{1}{1-e^{-x}} \quad (3.15)$$

Après chaque prédiction, les poids de pondération w_i seront ajustés selon l'équation suivante :

$$w_i \leftarrow w_i + \eta x_i (y - P(1)) \quad (3.16)$$

où η est la vitesse d'apprentissage et y désigne le symbole à coder.

Comme dans le cas du PPM, le CM utilise également un codage arithmétique après estimation des probabilités d'apparition. Comparé avec les deux autres codages par contexte, le CM bénéficie d'un modèle de prédiction plus flexible et par conséquent mieux adapté aux régularités locales des données. La considération des contextes de longueur différente et la modification des pondérations selon les performances des prédictions historiques font de cette méthode, l'un des algorithmes les plus performants qu'il existe.

Étant donnée la complexité de son modèle de prédiction, le CM est parfois critiqué pour son temps d'exécution relativement long. Or, dans notre étude, on cherche à estimer la complexité de Kolmogorov des séries financières. Le temps d'exécution des algorithmes n'est donc pas un critère à prendre en considération. La famille d'algorithmes sans perte PAQ, une des applications les plus efficaces de la méthode du CM, sera donc notre principal outil dans les analyses des rentabilités boursières.

Toujours dans l'optique d'améliorer le résultat de la compression, les techniques de transformation sont souvent utilisées pour “préparer” les données. Ces transformations réversibles, ayant pour vocation de regrouper les symboles identiques et de faire augmenter leur fréquence locale, sont présentées dans l'annexe L pour les lecteurs qui s'y intéressent.

Conclusion

Dans cette section, nous avons présenté 2 grandes catégories de méthodes de compression : i) les codages d'entropie, i.e., le codage de Huffman, le codage d'entropie et le codage par dictionnaire, ii) les codages par contexte.

Ces différentes techniques partagent un seul objectif : l'optimisation de l'espace pour le stockage des informations. Comme la compression est fondée sur des régularités, l'absence de régularités implique un taux de compression qui tend vers 0 quand la longueur des données est suffisamment importante. Le degré d'aléa d'une suite donnée peut ainsi être estimée par le meilleur taux de compression qu'on peut obtenir.

Or, chaque algorithme de compression a son "champ d'application", et n'est efficace qu'envers certains types de régularités bien spécifiques. Les régularités n'entrant pas dans leur "champ d'application" n'engendrent pas de compression valable. Par conséquent, le niveau d'aléa estimé par les algorithmes de compression présente un risque d'erreur. Dans nos études empiriques, nous contrôlons ce risque en employant une gamme d'algorithmes qui couvre toutes les catégories d'algorithmes présentées dans cette section, c'est-à-dire le codage de Huffman, le codage par dictionnaire et le codage par contexte.

L'objectif de ce chapitre est d'utiliser ces "détecteurs de régularités" pour estimer le niveau d'aléa des séries financières. Pour cela, le premier obstacle à franchir se trouve dans l'incompatibilité des algorithmes de compression avec les données financières. D'un côté, les rentabilités financières sont notées en nombres réels. D'un autre côté, les algorithmes de compression ne peuvent fonctionner que sur des séries discrètes (par exemple les suites des nombres entiers). Pour appliquer les algorithmes de compression aux séries financières avec sens, il convient de mettre en place un processus de transformation "nombres réels \rightarrow nombres entiers". Le processus que nous avons conçu à cet effet, fera l'objet de discussion dans la section suivante.

3.5 Détection des régularités financières avec les algorithmes de compression

Dans cette section, nous introduisons une méthode générale permettant de transformer les rentabilités financières en données informatiques adaptées aux algorithmes de compression.

3.5. Détection des régularités financières avec les algorithmes de compression

La présentation de cette méthode générale se déroulera en 3 étapes :

1. Illustrer avec des exemples concrets comment les algorithmes de compression détectent des régularités dans les séries de nombres entiers.
2. Introduire à travers un exemple plus élaboré, le processus de “discrétisation” permettant de généraliser la méthode algorithmique aux séries de nombres réels.
3. Enfin, comparer les algorithmes de compression avec les tests statistiques classiques, dans l’objectif de montrer avec des exemples concrets que certaines régularités détectables par les algorithmes de compression ne le sont pas toujours par des tests statistiques.

Afin d’illustrer le fonctionnement des “détecteurs de régularités”, nous allons étudier, dans un premier temps, une série de prix entiers qui contient des structures évidentes à révéler. Cette série a été engendrée par simulation pour illustrer la technique que nous proposons.

L’exemple à analyser est présenté dans la figure 3.8.

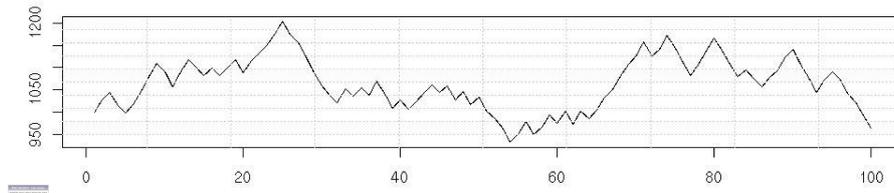


Fig. 3.8: Série de prix à analyser avec la méthode algorithmique.

Les 14 premières valeurs de cette chronique sont :

A: 1000, 1028, 1044, 1015, 998, 1017, 1048, 1079, 1110, 1090, 1058, 1089,
1117, 1100, . . .

Ici, une première régularité est manifeste : tous les prix semblent être regroupés au voisinage de 1000. On peut donc “effacer cette première régularité” en différenciant terme à terme la chronique, en appliquant l’équation $e_t = p_t - p_{t-1}$. La chronique des écarts de prix (e_t) est alors,

B: 28, 16, -29, -17, 19, 31, 31, 31, -20, -32, 31, 28, -17, -17, . . .

Nous souhaitons transformer cette suite numérique en base 10 en quelque chose se prêtant à la compression, c’est-à-dire une série en base 2. À cet effet, nous

3.5. Détection des régularités financières avec les algorithmes de compression

allons procéder une transformation algébrique très simple : comme tous les écarts de prix sont compris entre -32 et $+31$, la quantité “ $e_t + 32$ ” est toujours comprise entre 0 et 63 ; on passe de B à C en ajoutant à chacun de ses termes 32.

C : 60, 48, 3, 15, 51, 63, 63, 63, 12, 0, 63, 60, 15, 15, . . .

Puisque tout entier entre 0 et 63 peut être codé avec 6 bits d’information dans un système binaire, la suite C peut se transformer en une série de nombres binaires comme suit¹⁷ :

111100, 110000, 000011, 001111, 110011, 111111, 111111, 111111, 001100,
000000, 111111, 111100, 001111, 001111, . . .

ce qui conduit à la suite suivante, par concaténation :

D : 11110011000000001100111111001111111111111111111111110011
00000000111111111100001111001111. . .

Cette suite de 1 et de 0 nommée D est-elle compressible ? Une première “régularité” a été effacée quand nous sommes passés des prix aux variations des prix. La suite D occupe ainsi sensiblement moins d’espace que la suite A ¹⁸, cela correspond bien à une première compression où on utilise l’aspect séquentiel des valeurs qui tournent toutes autour de 1000.

Si D n’était pas compressible, nous serions arrivés au bout de notre recherche de régularités. Cependant, il se trouve que cette suite est compressible. Elle possède la régularité suivante : s’il y a un 0 (resp. 1) en position $2n$, alors il y a un 0 (resp. 1) en position $2n + 1$. Cette suite peut donc être compressée en ne conservant que les éléments en position paire, on obtient donc la suite D' présenté ci-dessous qui n’occupe que 42 bits.

D' : 110100001011101111111111010000111110011011. . .

17. Par exemple, le premier terme “60” en base 10 se note “111100” en base 2. En effet, $1 \times 2^5 + 1 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 = 60$

18. Chaque terme de la suite A est codé sur 4 octets. 56 octets ($56 \times 8 = 448$ bits) sont donc utilisés pour les 14 premières valeurs de la suite. En revanche, la suite D qui correspond aux variations de prix n’occupe que $6 \times 14 = 84$ bits.

À nouveau, nous pouvons croire être arrivé à une suite parfaitement aléatoire dans laquelle aucun algorithme de compression ne trouve plus de régularité. Pourtant, la suite D' a été engendrée par le générateur pseudo-aléatoire du langage de programmation “R” et donc, en théorie, pourrait se ramener à sa graine. Cependant, cette dernière étape ne peut être envisagée sérieusement en l'état actuel du développement des algorithmes de compression.

On voit ainsi comment se cachait, derrière le désordre apparent de la chronique des prix A , trois régularités importantes¹⁹ et au bout du compte un processus purement déterministe (l'algorithme de tirage de nombres pseudo-aléatoires).

L'exemple donné reste une illustration simplifiée de la logique “régularité-compression”. En étudiant directement une suite de nombres entiers, nous espérons avoir mis en évidence le mode de fonctionnement des algorithmes de compression sur les séries de prix.

Or, étant donnée la façon dont les rentabilités financières sont calculées ($r_t = \log(p_t) - \log(p_{t-1})$), les séries de rendements successifs, souvent exprimées en nombres réels, ne sont pas directement reconnaissables par les algorithmes de compression. Un processus de prétraitement est donc indispensable pour une application générale des outils informatiques. Ce processus sera introduit dans la section suivante qui aura pour objectif d'employer les algorithmes de compression²⁰ dans les détections des régularités.

3.5.1 *Processus de discrétisation : un prétraitement indispensable aux rentabilités financières*

L'enjeu de cette section est de montrer comment les algorithmes de compression peuvent nous permettre, au fil d'un certain nombre d'étapes, de formuler un jugement sur le degré d'aléa d'une série de rentabilités.

Comme nous l'avons évoqué à plusieurs reprises, l'application des algorithmes de compression aux séries financières nécessite un processus de discrétisation qui consiste à transformer les séries de nombres réels en suites de nombres entiers.

La série construite pour illustrer le processus de discrétisation mis en place, nommée $\sigma_0, \sigma_1, \dots, \sigma_n$, est obtenue à l'aide du langage “R”. Elle est constituée de 32000 réalisations d'une variable aléatoire distribuée selon une loi $N(0, 1)$ permettant de créer une chronique ressemblant à une série de cours boursiers. (cf.

19. Les trois régularités sont respectivement : valeur des prix dans un intervalle limité, structure particulière des nombres binaires et suite pseudo-aléatoire engendrée par une graine.

20. Il s'agit ici des implémentations des algorithmes de Huffman, Gzip, et PAQ8o8.

3.5. Détection des régularités financières avec les algorithmes de compression

figure 3.9 : en haut à gauche, l’histogramme des “rentabilités simulées”, au milieu, les cours simulés à partir des rentabilités artificielles avec une valeur initiale de 100). Les tests statistiques sur cette série ne dévoilent aucune structure spécifique

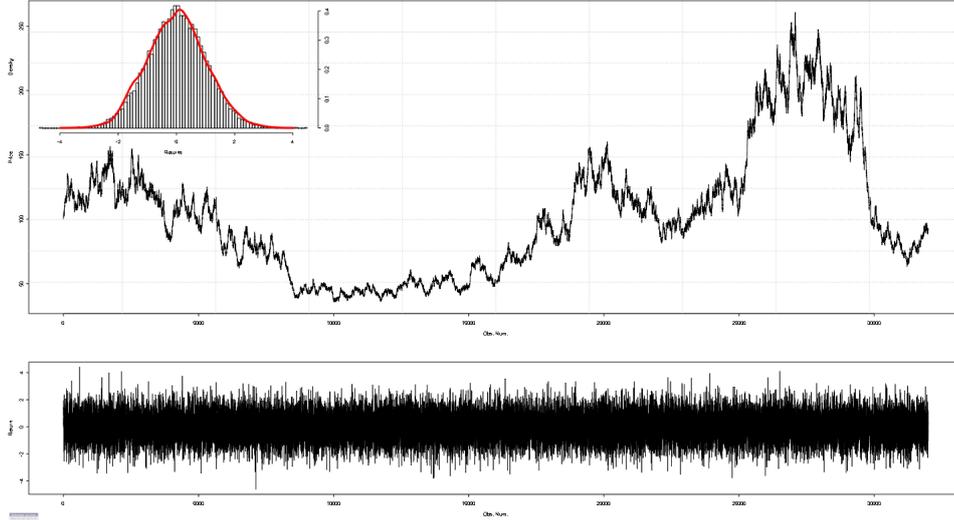


Fig. 3.9: Séries de cours et de rendements artificiels simulées à partir d’une variable aléatoire *i.i.d.N*

(les détails des tests sont donnés dans l’annexe I).

Comme nous l’avons évoqué précédemment, pour appliquer les algorithmes de compression, il faut transformer cette série en une suite de nombres entiers.

Notre première idée consiste à diviser les 32000 valeurs de rentabilités ordonnées (de la plus petite à la plus grande) en 256 intervalles de taille identique. Et à chaque intervalle, nous pouvons associer un nombre entier compris entre 1 et 256. Ainsi, chacun des 32000 termes observés peut être représenté par le nombre entier associé à l’intervalle qui le contient. En termes mathématiques, ce principe peut être décrit comme suit :

Calculons d’abord la taille de chaque intervalle, e selon l’équation (3.17) :

$$e = (M - m)/256 \quad (3.17)$$

où M désigne la valeur maximale observée dans $\sigma_0, \sigma_1, \dots, \sigma_n$ et m la minimale. À chaque terme, x , de la série $\sigma_0, \sigma_1, \dots, \sigma_n$ serait alors associé l’entier k tel que :

$$x \in [m + (k - 1) \times e, m + k \times e[\quad (3.18)$$

soit le numéro de l’intervalle auquel x appartiendrait parmi les 256 possibles. Dans l’équation (3.18), k est codé sur 8 bits d’information ($2^8 = 256$). On obtiendrait

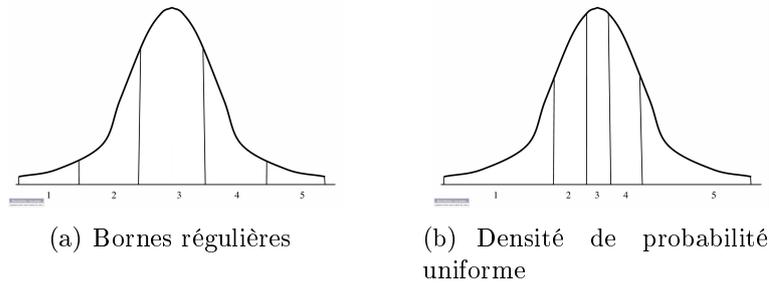


Fig. 3.10: Discretiser pour uniformiser : comment placer les bornes ?

alors un série de bornes régulières sur l’axe des abscisses sous la courbe de la loi normale, ainsi que représenté la figure 3.10(a). Bien évidemment, en utilisant cette première méthode de discrétisation, le fait que $\sigma_0, \sigma_1, \dots, \sigma_n$ suive une loi Normale se repérera dans les valeurs de k après transformation. Par exemple, il y aura beaucoup plus de k proches de 128 que de 1 ou 256. Un algorithme de compression détecterait cette régularité, et en l’exploitant, réussirait à compresser la suite des nombres entiers sans difficulté.

Cependant, la compression ainsi obtenue ne délivre aucune information supplémentaire concernant la nature de la suite d’origine, car sa loi de distribution est connue avant tout test de compression. Pour que les algorithmes de compression aillent plus loin que la loi normale, il faudra “effacer” cette loi et rendre uniformément distribuée la série après discrétisation. Ce processus nommé “uniformisation” est présenté dans la section suivante.

3.5.2 Effacement des régularités à l’aide des transformations algébriques

Toute loi de distribution non uniforme constitue une “structure massive” qui empêche les algorithmes de compression de révéler de nouvelles régularités structurelles plus “discrètes”. Nous proposons donc un processus d’uniformisation qui “efface” ces lois non uniformes suivies par les rentabilités financières.

L’illustration de ce processus se fera toujours avec la série présentée dans la figure 3.9. Après ce processus d’uniformisation, nous allons employer aussi bien les tests statistiques que les algorithmes de compression pour mesurer la quantité d’aléa dans la série $\sigma_0, \sigma_1, \dots, \sigma_n$. Comme celle-ci a été engendrée de manière indépendante, si la méthode d’uniformisation efface correctement la loi normale, les algorithmes de compression doivent arriver à la même conclusion que les tests statistiques : la suite uniformisée est totalement incompressible, elle est donc aléatoire.

Commençons toujours par la série de nombres réels présentée dans la figure 3.9. Cette fois-ci, au lieu de créer les intervalles de taille identique comme nous l'avons fait dans la section précédente, nous allons calculer les $borne(i)$ pour faire en sorte que l'aire sous la courbe de distribution soit de surface équivalente.

L'idée est de déterminer une série des valeurs croissantes $borne(1)$, $borne(2)$, \dots , $borne(255)$ de telle manière qu'un tirage selon la loi $N(0, 1)$ ait la même probabilité de tomber dans chacun des 256 intervalles définis par les $borne(i)$. (cf. figure 3.10(b)). Les intervalles ainsi calculés n'ont pas toujours la même longueur : ils sont d'autant plus étroits qu'ils sont proches de la valeur centrale 0.

Remarquons que cette méthode de discrétisation ne s'applique pas uniquement aux séries normales. Même sans connaître la distribution des données initiales (ce qui est souvent le cas des rentabilités boursières), nous pouvons toujours fixer les $bornes[i]$ à l'aide de la distribution empirique issue des données : il suffit de classer les valeurs observées par ordre croissant et les séparer en 256 sous-ensembles ayant des effectifs aussi égaux que possible. Les 255 bornes à identifier sont donc les frontières entre les 256 sous-ensembles ainsi obtenus.

La chronique des rendements après uniformisation est présentée dans la figure 3.11. Chaque point de la figure désigne la valeur d'un rendement σ_i sur une échelle comprise entre 0 et 255. Le "semis" ainsi obtenu présente un aspect très homogène, sans zone particulièrement dense ou désertée²¹. Un rapide examen visuel confirme ainsi que l'uniformisation semble effective.

Comme chaque entier entre 0 et 255 correspond à un caractère `ascii`, la série de rentabilités uniformisée est en réalité un texte contenant 32000 caractères `ascii`. Ce texte peut-il être compressé ? En d'autres termes, présente-t-il des régularités invisibles à l'oeil mais détectable par les algorithmes de compression²² ?

Lorsqu'on applique les algorithmes de compression sur le texte obtenu à partir des caractères `ascii` (cf. figure 3.12), ceux-ci s'avèrent complètement inefficaces (cf. tableau 3.7).

Cet exemple a mis en évidence l'efficacité du processus d'uniformisation : comme les algorithmes de compression ne peuvent pas réduire la longueur du texte après uniformisation (cf. 3.12), nous inférons que la complexité de Kolmogorov de la suite des rentabilités artificielles ici étudiées est égale à sa longueur initiale.

21. On verra dans les sections prochaines que le cas contraire peut se produire si la chronique initiale n'est pas aléatoire.

22. D'un point de vue pratique on transforme la chronique des rentabilités en un texte en associant à chaque rentabilité observée aux différentes dates $t \in [1, 32000]$ l'entier lui correspondant.

3.5. Détection des régularités financières avec les algorithmes de compression

Celle-ci peut donc être considérée comme aléatoire.

Différents tests statistiques sont également appliqués à la série $\sigma_0, \sigma_1, \dots, \sigma_n$, ceux-ci ne dévoilent évidemment aucune structure spécifique, ce dont on pourra se rendre compte en consultant les tableaux 3.8, 3.9, 3.10.

Tab. 3.8: Tests d'autocorrélation pour la série *i.i.d.N*

Série	$\chi - deux$	deg. lib.	<i>p-value</i>
$\sigma_0, \sigma_1, \dots, \sigma_n$	0.1838	1	0.668
	36.9157	36	0.4264

Interprétation : *les termes de la série $\sigma_0, \sigma_1, \dots, \sigma_n$ sont uncorrélés.*

Tab. 3.9: Test BDS pour la série *i.i.d.N*, $m = \{2, 3\}$

ε	0.5012	1.002	1.5035	2.0047
$m = 2$	-0.2082	-0.2987	-0.5232	-0.7221
<i>p-value</i>	0.8351	0.7651	0.6009	0.4702
$m = 3$	0.8351	0.7651	0.6009	0.4702
<i>p-value</i>	0.9503	0.9803	0.9344	0.8600

Interprétation : *les termes de la série $\sigma_0, \sigma_1, \dots, \sigma_n$ sont i.i.d..* Ici, ε et m sont les paramètres classiques du test BDS que nous avons présenté en détail dans l'annexe F. On constate dans ce tableau que, sous hypothèse nulle "la série examinée est *i.i.d.*", les *p-values* obtenues sont toujours $> 20\%$, quelle que soit la valeur des paramètres retenue.

Tab. 3.10: Tests de stationnarité pour la série *i.i.d.N*

Test	Val. statistique	ordre	<i>p-value</i>
ADF	-32.7479***	31	0.01
Phillips-Perron (PP)	-178.52949***	16	0.01

Interprétation : *La série $\sigma_0, \sigma_1, \dots, \sigma_n$ est stationnaire.*

Ces résultats indiquent une certaine cohérence entre la méthode algorithmique et statistique face aux données aléatoires. Or, cela n'implique pas l'équivalence des deux approches envers des suites régulières. Sur le plan théorique, les tests statistiques et les algorithmes de compression sont développés dans deux disciplines relativement indépendantes, ils n'ont pas été conçus pour révéler les mêmes types de régularités. Il sera donc tout à fait logique que certaines régularités soient détectables uniquement par l'une des deux approches utilisées.

En effet, pour illustrer l'intérêt des algorithmes de compression par rapport aux tests statistiques, nous allons montrer dans la prochaine section que *cer-*

taines régularités détectables par les algorithmes de compression ne le sont pas par les tests statistiques.

3.5.3 Algorithmes de compression vs. tests statistiques

Certaines chroniques ordonnées peuvent paraître aléatoires aux yeux des tests statistiques traditionnels. L'idée que de telles structures existent est bien connue dans la théorie des systèmes dynamiques, où des équations déterministes, qui forment par ailleurs de telles "structures lourdes", engendrent des séries de données qui paraissent imprévisibles (cf. phénomène chaotique).

Dans cette section, nous allons montrer successivement, sur la base de chroniques artificielles dont nous contrôlons parfaitement la structure, comment certaines régularités demeurant invisibles aux méthodes statistiques peuvent être révélées par les algorithmes de compression.

Commençons ces opérations sur une chronique artificielle dans laquelle nous avons pris le soin de cacher une structure régulière importante. Un point explicatif est ici nécessaire pour clarifier le procédé par lequel nous avons "dissimulé" une telle structure dans la chronique.

Plusieurs types de structures sont envisageables. Certaines régularités sont repérables sans difficulté par les tests statistiques standards (processus autorégressifs ou processus non stationnaires par exemple). Ce ne sont évidemment pas les structures qui nous intéressent. Nous cherchons donc les régularités que nulle méthode statistique classique ne détectera, du moins aisément.

À cet effet, nous créons tout d'abord une séquence de nombres entiers compris entre 0 et 255, nommée *texte*. Au départ, *texte* est obtenue en effectuant 32000 tirages aléatoires uniformes, durant lesquels chaque entier a la même probabilité d'être tiré, c'est-à-dire $\frac{1}{256}$.

La séquence *texte* sera soumise à diverses transformations dans l'objectif d'y "implanter" des régularités indétectables par les tests statistiques. La séquence après transformations, est notée *texte'*, afin de la distinguer du texte uniforme non biaisé.

La structure que nous allons "implanter" dans *texte* pour obtenir *texte'* consiste en une alternance de 0 et de 1 sur le dernier (cas 1) ou les trois derniers (cas 2) digits de chaque terme du *texte* écrit en base 2.

3.5. Détection des régularités financières avec les algorithmes de compression

Par exemple, en codant les nombres entiers du *texte* sur 8 bits binaires, nous devons obtenir, dans le cas 1, un *texte'* du type

00000001, 000110110, 11101001, 100001110, 10000111,...

Cette régularité correspond à une “alternance de parité” puisque les nombres binaires terminant par 1 sont les nombres pairs et ceux terminant par 0 sont impairs. Alors, dans le cas 2 nous souhaitons obtenir des séries du type :

01010000, 11101001, 01101010, 01101011, 11100100, 10100101, 11010110,
11100111, 00011000,...

Après ces transformations, nous disposons d’une séquence biaisée de 32000 termes tirés de façon uniforme entre $[0, 255]$.

La seconde étape consiste alors à tirer 32000 nombres réels de la variable aléatoire *i.i.d.* $N(0, 1)$ et de calculer 256 bornes ($borne_0, borne_1, \dots, borne_{255}$) de façon à obtenir une densité de probabilité uniforme entre chacune d’elles (ce qui correspond à la méthode d’uniformisation représentée par la figure 3.10(b) page 135).

Nous associons ensuite à chaque terme de *texte'*, noté α_i , un nombre réel tiré de façon uniforme de l’intervalle $[borne_{\alpha_i}, borne_{\alpha_i+1}]$. Nous obtenons ainsi une chronique notée *chron* qui, par construction, vérifiera les deux propriétés suivantes :

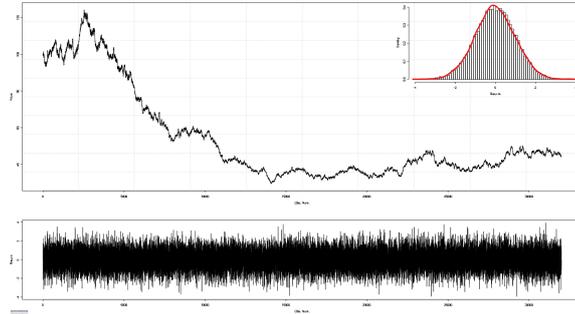
1. globalement, les valeurs de *chron* suivront une loi normale.
2. la suite des numéros des intervalles ($[borne_i, borne_{i+1}]$) dans lesquels les termes de *chron* étaient tirés, sera exactement *texte'*.

La structure cachée dans *chron* sera-elle détectée par les tests statistiques ? Quelle sera la performance des algorithmes de compression ? C’est ce que nous allons maintenant étudier.

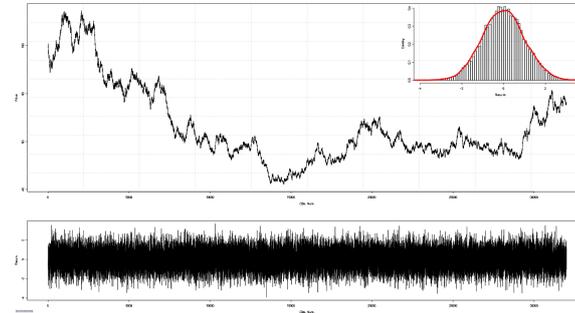
Afin de bien illustrer les données, nous transformons d’abord *chron*, la pseudo-chronique de rentabilités, en une série de cours selon l’équation : $p_0 = 100, p_t = p_{t-1}e^{chron[t]}$. La série de prix ainsi obtenue est présentée dans la figure 3.13. La figure 3.13(a) correspond au cas où une alternance de parité est cachée dans la chronique des rendements (cas 1) et la figure 3.13(b) représente le cas 2 (structure dans les 3 derniers digits).

Les tests statistiques sur ces séries biaisées ne dévoilent aucune structure spécifique (cf. tableaux 3.11, 3.12, 3.13) **et sont donc en échec.**

3.5. Détection des régularités financières avec les algorithmes de compression



(a) Cas 1



(b) Cas 2

Fig. 3.13: Séries simulées cachant deux structures implantées

Tab. 3.11: Tests d'auto-corrélation pour *chron*

cas1			cas2		
χ -deux	deg. lib.	<i>p-value</i>	χ -deux	deg. lib.	<i>p-value</i>
0.0096	1	0.9219	1.1169	1	0.2906
29.6655	36	0.7629	45.4802	36	0.1337

Interprétation pour cas 1 : *les termes de la série sont indépendants.*

Interprétation pour cas 2 : *les termes de la série sont indépendants.*

Tab. 3.12: Test BDS pour *chron*, $m = \{2, 3\}$

cas 1					cas 2				
ε	0.5006	1.0012	1.5019	2.0025	ε	0.4988	0.9976	1.4964	1.9952
$m = 2$	0.1121	0.1790	0.3377	0.4318	$m = 2$	0.1886	0.1186	0.0627	0.1207
<i>P-V</i>	0.9108	0.8579	0.7356	0.6659	<i>P-V</i>	0.8504	0.9056	0.9500	0.9039
$m = 3$	0.1329	0.2233	0.3578	0.4662	$m = 3$	-0.0974	-0.0617	0.0336	0.2153
<i>P-V</i>	0.8943	0.8233	0.7205	0.6411	<i>P-V</i>	0.9224	0.9508	0.9732	0.8295

Interprétation pour cas 1 : *les termes de la série sont i.i.d.*

Interprétation pour cas 2 : *les termes de la série sont i.i.d.*

3.5. Détection des régularités financières avec les algorithmes de compression

Tab. 3.13: Tests de racine unitaire pour *chron*

Test	cas 1			cas 2		
	V.stat	ordre	<i>P-V</i>	V.stat	ordre	<i>P-V</i>
ADF	-31.5598***	31	0.01	-31.2825***	31	0.01
PP	-178.797***	16	0.01	-179.9449***	16	0.01

Interprétation pour cas 1 : *la série est non stationnaire.*

Interprétation pour cas 2 : *la série est non stationnaire.*

Après avoir constaté la performance des tests statistiques au regard de la séquence *chron*, nous appliquons sur les deux chroniques biaisées (cas 1 et cas 2) le processus d'uniformisation que nous avons introduit, dans la section précédente. Comme l'uniformisation est en réalité, l'inverse de la transformation "*texte'* → *chron*", la série *chron* uniformisée sera exactement le texte du départ *texte'*.

La régularité dissimulée dans le cas 1 implique en théorie une compression de 12.5%. Étant donnée l'alternance de parité, chaque terme de *texte'* codé sur 8 bits d'information n'en nécessite que 7 : en observant le premier terme de la série, les derniers digits de tous les termes suivants seront connus d'avance. On gagne en théorie 1 bit sur 8 en exploitant cette régularité, ce qui donne un taux de compression théorique de 1/8.

En suivant cette même logique, le taux de compression théorique dans le deuxième cas, se relève à $3/8 = 37,5\%$, car cette fois-ci, les 3 derniers digits de chaque terme sont parfaitement prévisibles.

Pour vérifier ces conclusions théoriques, nous avons appliqué des algorithmes de compression sur les séries *chron* uniformisées. Les résultats de compression sont exposés dans le tableau 3.14. On constate ici que les trois algorithmes utilisés réalisent ***tous une belle performance de compression***, notamment PAQ8o8 qui a atteint les taux de compression très proches de leur valeur théorique. Ce résultat illustre bien l'efficacité des algorithmes dans la recherche des régularités.

Cet exemple illustre l'intérêt de l'approche algorithmique dans les recherches de régularités : les algorithmes de compression sont parfois capables de révéler les régularités indétectables par les tests statistiques traditionnels.

Or, dans certains cas de figure, les algorithmes de compression *actuellement disponibles* sont également mis en échec, comme nous allons le montrer dans la section suivante.

3.5. Détection des régularités financières avec les algorithmes de compression

Tab. 3.14: Tests de compression pour *texte'*

Algorithme	cas 1		cas 2	
	taille d'origine	taux réalisé	taille d'origine	taux réalisé
	32000	0%	32000	0%
Huffman	31235	2.39%	23079	27.88%
Gzip	31322	2.12%	23160	27.63%
PAQ8o8	28296	11.58%	20974	34.46%

Interprétation pour cas 1 : *On peut compresser la série uniformisée*

Interprétation pour cas 2 : *On peut compresser la série uniformisée*

3.5.4 Limites pratiques de la méthode algorithmique : le cas des décimales de π

Le nombre π fait parti des plus étudiés en mathématique, avec peut-être le nombre d'Euler et le nombre Ω de Chaitin. Il existe de très nombreuses façons de le calculer. Les deux formules ci-dessous donnent des moyens d'inégale simplicité pour calculer un grand nombre de ses décimales.

– Formule de Leibnitz-Madavar :

$$4 \times \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \quad (3.19)$$

– Formule en développement :

$$\pi = \lim_{n \rightarrow \infty} \sqrt{6 \times \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \dots + \frac{1}{n^2}\right)} \quad (3.20)$$

Comme nous l'avons fait pour l'alternance de parité, on peut cacher π dans une série de cours en trois étapes :

- Chaque décimale de π est codée en base 2 sur 4 bits. Par exemple, les 4 premières décimales (1415) formeront la séquence 0001, 0100, 0001, 0101. En s'intéressant aux 50000 premières décimales de π , on obtient 200000 bits d'information.
- On reconsidère la séquence binaire obtenue par octet. Par exemple, pour les 4 premières décimales on obtient deux octets, 0001010000010101. En base 10, chacun de ces octets correspond à un entier compris entre 0 et 255 (les deux octets donnés en exemple correspondent aux nombres 20 et 21).
- Il reste à associer à chacun de ces entiers une valeur tirée de façon uniforme entre les deux bornes de son intervalle correspondant. Les bornes ont été calculées grâce à la distribution $N(0,1)$, comme nous l'avons présenté à la section 3.5.2 (page 135). On obtient ainsi une pseudo-chronique de rentabilités qui elle même permet d'engendrer la pseudo chronique de cours représentée dans la figure 3.14.

3.5. Détection des régularités financières avec les algorithmes de compression

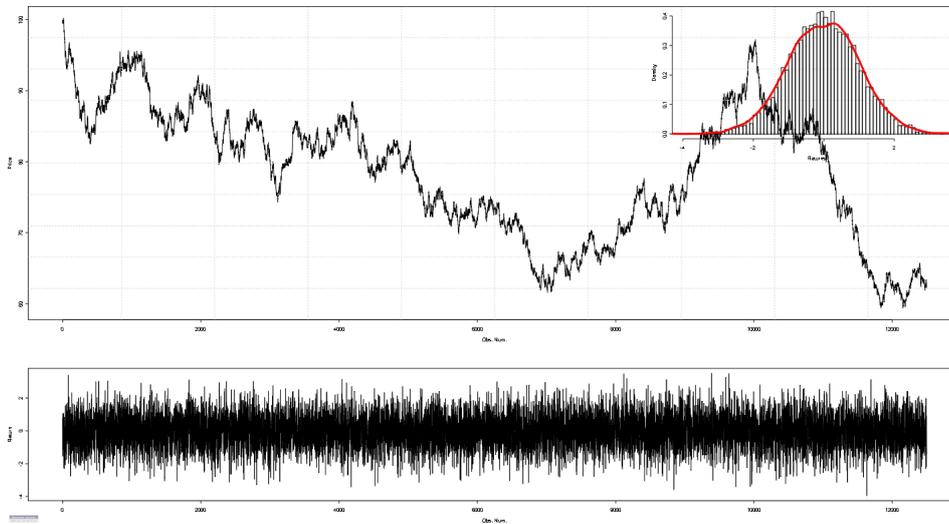


Fig. 3.14: Pseudo-chronique financière engendrée par la séquence des décimales de π . En haut les cours, en haut à droite, la distribution des pseudos-rentabilités, en bas leur chronique.

La série de π n'est pas compressible après le processus précédemment expliqué, car aucun algorithme actuellement disponible ne recherche π dans une séquence de digits à notre connaissance. Pourtant, l'existence de formules pour calculer π montre bien que sa complexité se réduit à peu de chose, en termes de théorie algorithmique de l'information. Cette conclusion est aussi confirmée par nos tests de compression dont les résultats sont présentés dans le tableau 3.15. Cet exemple

Tab. 3.15: Tests de compression : π

Algorithme	taille du fichier	taux de compression
	12500	100%
Gzip	12566	- 0.528%
Huffman	12955	-3.64%
PAQ8o8	12587	-0.70%

Interprétation : *On ne peut pas compresser la série issue de π*

illustre clairement qu'il est toujours possible d'envisager qu'une régularité soit cachée derrière des manifestations aléatoires. Il est dans la pratique impossible de concevoir un algorithme de compression ultime qui exploite toutes les régularités pouvant exister dans une série de données binaires.

On constate toutefois que les tests statistiques classiques ne font pas mieux sur la série de π : aucun d'eux ne permet de rejeter l'hypothèse nulle associée à

3.5. Détection des régularités financières avec les algorithmes de compression

l'absence de régularité (cf. tableaux 3.16,3.17,3.18).

Tab. 3.16: Tests d'autocorrélation sur la série de π

$\chi - deux$	deg. lib.	$p-value$
2.7339	1	0.09824

Interprétation : *les termes de la série sont indépendants*

Tab. 3.17: Test BDS pour la série construite sur π , $m = \{2, 3\}$

ε	0.5023	1.0046	1.5069	2.0092
$m = 2$	-0.0895	0.0468	0.0395	0.0129
$p-value$	0.9287	0.9627	0.9685	0.9897
$m = 3$	-0.4005	-0.1755	-0.2780	-0.3208
$p-value$	0.6888	0.8607	0.7810	0.7483

Interprétation : *On ne peut pas rejeter l'hypothèse nulle, les termes de la série sont donc i.i.d.*

Tab. 3.18: Tests de runs pour la série construite sur π

Statistique W	$p-value$
-1.368	0.1713

Interprétation : *On ne peut rejeter H_0 , les termes de la série sont indépendants*

Il n'en demeure pas moins que cette limite est circonstancielle et temporaire, et qu'elle pourrait **théoriquement** être dépassée avec l'invention des nouveaux algorithmes de compression.

Ces exemples artificiels développés, nous avons montré en quoi les méthodes algorithmiques sont susceptibles de détecter des structures régulières que les méthodes statistiques ne semblent pas capables de détecter. Nous avons illustré également les limites actuelles de cette nouvelle méthode.

3.5.5 Une discussion méthodologique

Avant de commencer des applications sur données réelles, nous estimons nécessaire de récapituler la méthodologie proposée qui, jusqu'à présent, n'a fait que l'objet d'illustrations dans des exemples concrets. Dans cette récapitulation, nous défendons une méthode algorithmique qui se compose de 4 étapes successives :

Étape I : Convertir les séries de cours en suites de rentabilités. Les séries de cours financières présentent toujours un aspect séquentiel, elles sont donc fortement autocorrélées et compressibles. Sans la conversion "prix

Tab. 3.19: Tests de racine unitaire pour la série construite sur π

Test	Val. statistique	ordre	<i>p-value</i>
ADF	-23.3799***	23	0.01
PP	-110.1364***	13	0.01

Interprétation : *On peut rejeter H_0 , la série est non stationnaire*

→ rentabilité”, cette propriété des prix rendra difficile l’interprétation des résultats empiriques.

Etape II : Transformer les rentabilités obtenues en nombres entiers.

Ce processus nommé “discrétisation” est indispensable dans les études de données financières. Souvent, les rentabilités successives calculées selon l’équation $\sigma_t = \log(p_t) - \log(p_{t-1})$ sont notées en nombres réels, alors que les algorithmes de compression traitent uniquement les données discrètes. L’application des outils informatiques est donc dénuée de sens sans la discrétisation des rentabilités.

De plus, la représentation des rentabilités en nombres réels suscite souvent une précision illusoire, au sens où seul un certain nombre de décimales est intéressant sur le plan pratique. Par exemple, il y a peu de chances qu’on s’intéresse au 18ième chiffre après la virgule d’une rentabilité boursière.

Nous pouvons donc transformer ces rentabilités inutilement précises en nombres entiers (par exemple compris entre 0 et 255), avec chacun de ces nombre codant un intervalle encadrant les rentabilités correspondantes.

Bien entendu, le nombre d’entiers retenus à cet effet est fixé aux puissances de 2 (dans notre exemple il est fixé à $2^5 = 256$), de façon à pouvoir les recoder en base 2 avec une occupation d’espace optimale. Par exemple, en utilisant 5 entiers dans la discrétisation, chacun d’eux doit être codé sur 3 bits d’information. Or, avec 3 bits d’information on aurait pu coder $2^3 = 8$ entiers différents, il y a donc une occupation sous optimale des espaces, causée par le choix du nombre d’entiers à utiliser dans le codage.

En écrivant les numéros d’intervalles en base 2, les suites de rentabilités deviendront des séquences de 1 et de 0, reconnaissables par des algorithmes de compression.

À ce niveau, deux interrogations nous paraissent naturelles :

1. Pourquoi ne pas essayer de compresser directement la chronique des rentabilités en base 10 ?

La réponse est la suivante : dans un ordinateur, toutes les informations sont stockées en base 2, chaque digít en base 10 (qui par définition varie entre 0 et 9) occupe 8 bits (1 octet). Par conséquent, en écrivant les rentabilités en base 10, on constitue un fichier qui n’utilise que 10 nombres parmi les 256 qu’on peut coder dans un octet²³. Donc, quelle

23. En système binaire, chaque bit peut être soit 0 soit 1, donc avec 1 octet (8 bits) on peut

que soit la suite de rentabilités (aléatoire ou non), une compression due au stockage sous optimal des digits est toujours possible.

Pour résumer, les algorithmes de compression travaillent toujours sur des fichiers binaires. Il faut donc “organiser” les données de façon à pouvoir apprécier directement les taux de compression obtenus, sans biais de représentation (du type “10 nombres utilisés parmi 256 possibles”), ce qui impose l’usage de la base 2. Ici, il s’agit d’une transcription convenable permettant l’interprétation directe des résultats de compression.

2. En transformant les rentabilités réelles en un ensemble limité de nombres entiers, ne perd-on pas une partie des informations initiales ? De toute manière, en associant le même entier, disons 134, aux deux rentabilités proches mais différentes (par exemple “1,25% et 1,26%”), on diminue sans aucun doute le niveau de précision.

Ici, la réponse est la suivante. Il faut bien choisir le nombre d’entiers à utiliser dans le processus de discrétisation, de manière à conserver un degré de précision adéquat : ni illusoirement précis, ni grossièrement imprécis.

Ici, on procède comme en photographie numérique, il faut choisir la granularité de manière à ne perdre rien d’important dans la prise de vue. Par exemple, s’il est jugé nécessaire de retenir 4 chiffres après la virgule pour une série des rentabilités dont l’écart entre le maximum et le minimum est de 8%, on retiendra $2^{13} = 8192$ nombres entiers dans le processus de discrétisation.

Étape III : Si la série de rentabilités contient des régularités connues, il faut concevoir des transformations algébriques pour les “effacer” successivement. Par exemple, le processus d’uniformisation illustré par la figure 3.10(b) à la page 135 est conçu pour “effacer” les lois de distribution non uniformes éventuellement suivies par les variations de prix.

D’autres transformations pourraient être envisagées, si les données représentent d’autres structures connues. Par exemple, dans la prochaine section, nous allons introduire une transformation algébrique visant à effacer le phénomène des “bouffées de volatilité”.

Après l’effacement des structures connues, chaque nombre entier est représenté par son code `ascii` correspondant. La suite de rentabilités réelle se convertit ainsi en un fichier de texte prêt à la compression.

Étape IV : Les algorithmes de compression sont utilisés pour révéler de nouvelles régularités. Dans ce chapitre, nous utilisons 3 algorithmes de compression qui représentent chacun une méthode d’encodage différente. “Gzip”²⁴ est un des algorithmes les plus performants dans la famille des codages par dictionnaire. “Huffman” applique scrupuleusement le

coder au maximum $2^8 = 256$ chiffres différents.

24. Disponible sur le site <http://www.gzip.org/index-f.html>

codage de Huffman présenté dans la section 3.4.1 (page 118). “PAQ8o8”²⁵ est un des algorithmes par contexte les plus efficaces en termes de taux de compression. Ces algorithmes utilisés couvrent les trois grandes catégories de compression sans perte. Leur combinaison doit pouvoir détecter la plupart des régularités actuellement exploitables par les algorithmes de compression.

Le choix de ces 3 outils par rapport aux algorithmes de la même famille est entièrement fondé sur le critère du taux de compression. En effet, pour chaque méthode de compression, différents modèles sont développés pour arbitrer entre le taux de compression et le temps d’exécution. En fonction de leurs objectifs, les utilisateurs exigent différentes relations entre ces deux critères qui évoluent souvent dans des directions opposées. Pour certaines activités, notamment les sites d’Internet qui compressent leurs données pour une transmission rapide, l’importance est d’obtenir un taux satisfaisant dans un délai raisonnable, mais non le meilleur taux de compression.

Or, l’objectif de notre étude étant de mesurer la quantité d’aléa dans des séries boursières, l’obtention du meilleur taux de compression est une priorité absolue. Par rapport aux autres algorithmes de la même famille, “Gzip”, “Huffman” et “PAQ8o8” consacrent leur rapidité au taux de compression optimal. Ils sont donc choisis pour les recherches des régularités financières.

L’interprétation des résultats obtenus est fondée sur les deux principes suivants :

- La compressibilité de la série des entiers implique la présence des régularités. Une suite de rentabilités est d’autant plus régulière que son taux de compression est élevé.
- L’incompressibilité de la suite s’interprète de manière plus compliquée :
 1. Soit la suite de rentabilités est réellement incompressible, donc aléatoire.
 2. Soit elle contient des régularités indétectables par les algorithmes utilisés, mais d’autres outils de compression sont capables de les exploiter. C’est par exemple le cas du nombre π .

Selon la théorie de la calculabilité, ces deux cas de figure ne peuvent pas être distingués de manière exacte. La complexité de Kolomogorov d’une suite donnée n’est pas calculable. Autrement dit, aucune machine de Turing ne peut associer à chaque suite s , la valeur exacte de sa complexité de Kolmogorov $K(s)$. Quel que soit le nombre d’algorithmes utilisés, ils n’exploiteront jamais toutes les régularités qu’il pourrait exister dans une série de rentabilités.

La méthode algorithmique proposée dans ce chapitre vise à définir explicitement une série d’algorithmes de compression notée C_0, C_1, \dots, C_n , et

25. Disponible sur le site <http://cs.fit.edu/~mmahoney/compression/>

3.5. Détection des régularités financières avec les algorithmes de compression

à estimer approximativement la complexité de chaque suite finie s selon le principe suivant :

$$\min_{i=0}^n T(C_i, s) = K(s) \quad (3.21)$$

où la fonction $T(C_i, s)$ calcule la longueur de s après la compression effectuée par C_i . Dans les exemples précédents, C_0, C_1, \dots, C_n contient trois algorithmes : Huffman, Gzip et PAQ808. D'une manière générale, plus C_0, C_1, \dots, C_n contient d'algorithmes, plus l'estimation de $K(s)$ sera proche de sa valeur théorique.

Les étapes présentées ci-dessus sont explicitement illustrées dans la figure 3.15, qui servira comme le guide schématisé des études empiriques dans les prochaines sections.

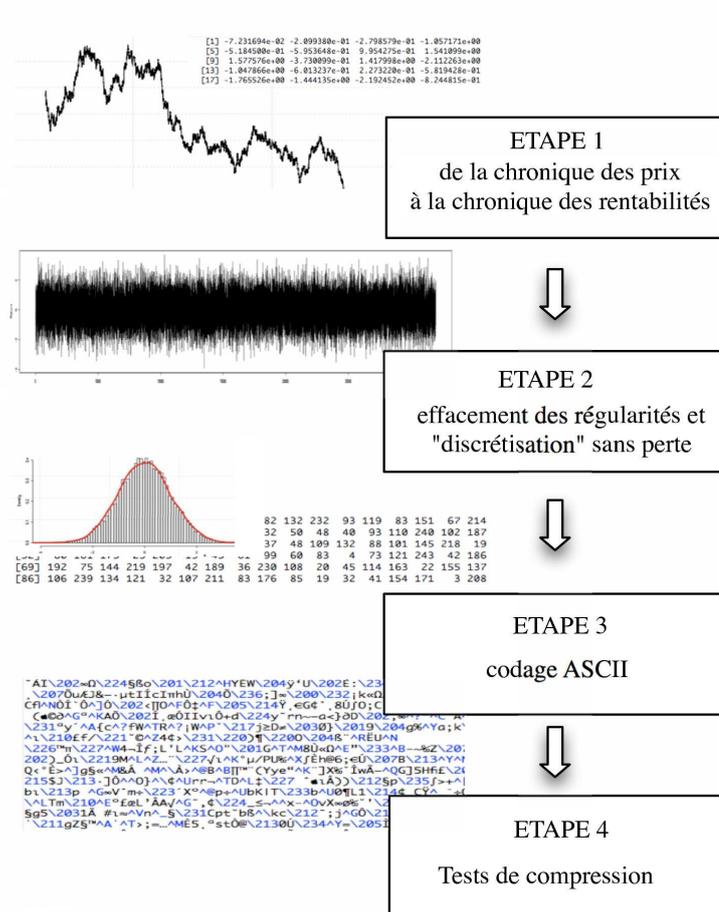


Fig. 3.15: Organisation générale des tests

3.6 Une première application sur les données réelles : le cas de l'indice Dow Jones

Désormais, les algorithmes de compression seront utilisés sur de vraies données financières dans l'objectif de mesurer le degré d'aléa de chacune des séries analysées. Pour commencer, la chronique longue du Dow Jones Industriel (1896-2005), observée sur une fréquence journalière, sera analysée comme une première mise en pratique de l'approche algorithmique. Chaque étape de cette étude sera illustrée avec tous les détails et figures nécessaires à sa compréhension.

Nous montrons à l'issue de cette analyse que les variations journalières de l'indice Dow Jones (désormais DJ) représentent un degré d'aléa élevé durant l'ensemble de son histoire. Après l'effacement successif des “faits stylisés” qui sont, on le rappelle, autant de “régularités”, la série de rentabilités est quasi-incompressible par les algorithmes utilisés.

3.6.1 Détection des faits stylisés

Après la transformation “*prix* \rightarrow *rendements*”, la série DJ obtenue contient 27423 rentabilités, comme présenté dans la figure 3.16. Après le processus de

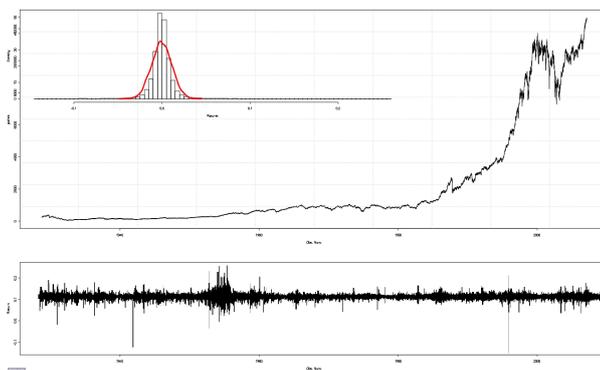


Fig. 3.16: Clôtures quotidiennes du DJ durant la période de 1928 à 2005.

discrétisation introduit dans la section 3.5.1 à la page 135, la chronique du DJ est transformée en une suite de nombres entiers qui suit une loi quasi-normale (cf. figure 3.17).

Les tests de compression sont conduits sur cette suite de nombres entiers. Les résultats obtenus sont exposés dans le tableau 3.20.

3.6. Une première application sur les données réelles : le cas de l'indice Dow Jones

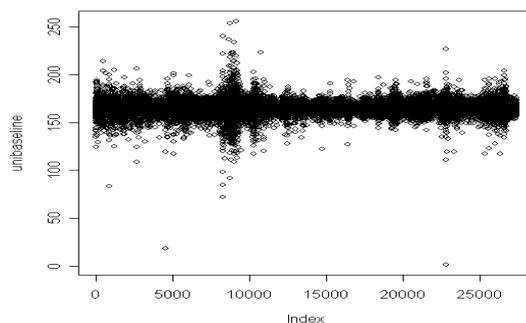


Fig. 3.17: Après discrétisation, la chronique journalière du DJ suit une loi quasi-normale.

Tab. 3.20: Tests de compression : variations de DJ avant l'uniformisation

Algorithme	taille du fichier	taux de compression
	27423	100%
Gzip	18038	34.22%
Huffman	18090	34.03%
PAQ8o8	18090	35.01%

Interprétation : *Sur les variations du DJ avant uniformisation, les algorithmes utilisés réalisent tous des taux de compression très élevés, car ils exploitent la loi de distribution non uniforme suivie par les rentabilités.*

3.6. Une première application sur les données réelles : le cas de l'indice Dow Jones

Il est clair qu'une structure régulière existe dans cette chronique : les rentabilités de l'indice sont distribuées approximativement selon une loi gaussienne. Cette régularité flagrante est signalée par les algorithmes utilisés en affichant des taux de compression considérables (cf. tableau 3.20).

Pour mesurer la quantité d'aléa *sans l'effet de la loi de distribution*, il faut effacer la régularité observée dans le tableau 3.20, en appliquant le processus d'uniformisation. Les bornes utilisées à cet effet sont calculées selon des observations empiriques, afin d'éviter toute hypothèse sur la loi de distribution des rentabilités boursières. Ce faisant, quelle que soit la loi statistique suivie par les rentabilités du DJ, nous pouvons toujours l'effacer en appliquant le processus d'uniformisation.

La figure 3.18 présente le “semis” des rentabilités du DJ après uniformisation. On s'aperçoit ici, et les tests de compression le vérifieront, que le processus d'uni-

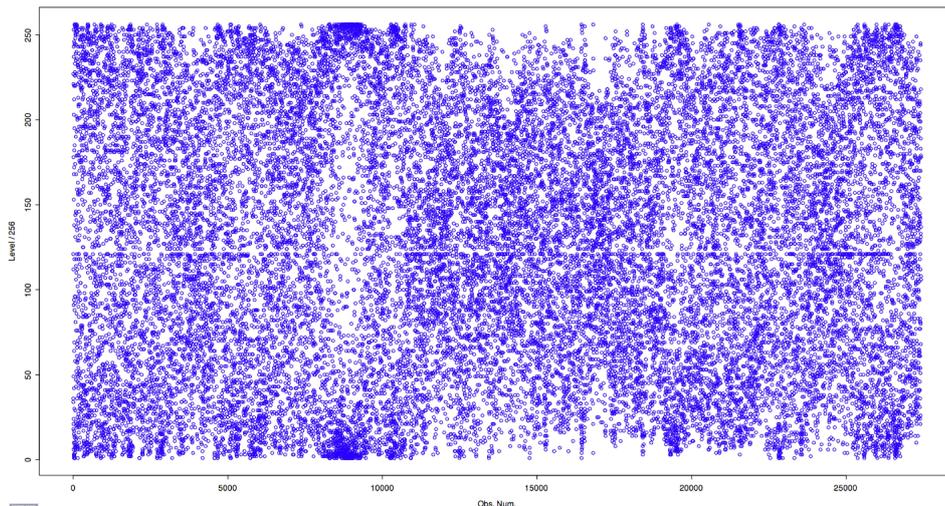


Fig. 3.18: Log-rentabilités du DJ après discrétisation et uniformisation

formisation mis en place ne produit pas de suite uniforme : le “semis” des valeurs n'est pas homogène et certaines zones semblent moins denses que d'autres. C'est par exemple le cas pour la zone comprise entre l'observation 9000 et 10000 autour de la valeur 128, et la zone 13000 – 20000 autour des valeurs extrêmes ²⁶.

Cette différence d'intensité est due au fait qu'il existe une autre régularité bien connue dans la chronique des rentabilités, et que nous ne sommes pas allés assez loin dans le processus d'effacement de structures : les volatilités des séries financières ne sont pas stables dans les temps. Les “bouffées de volatilité” sont

²⁶ C'est-à-dire les valeurs juste au-dessus de 0 en ordonnée ou juste en-dessous de 255

3.6. Une première application sur les données réelles : le cas de l'indice Dow Jones

les sources d'hétérogénéité de la figure 3.18. Par exemple, la zone 13000 – 20000 correspond à une période fortement volatile du marché, car les valeurs extrêmes y apparaissent plus souvent que les rentabilités de faibles envergures.

Les résultats de compression (cf. tableau 3.21) confirment la présence d'une telle régularité. Malgré les taux négatifs de Gzip et Huffman, l'algorithme par contexte, PAQ8o8 a pu exploiter le phénomène des bouffées de volatilité et réduire la taille du fichier d'origine. En constatant un faible taux de compression

Tab. 3.21: Tests de compression : Dow Jones après l'uniformisation

Algorithme	taille du fichier	taux de compression
	27423	100%
Gzip	27489	-0.24%
Huffman	27456	-0.12%
PAQ8o8	27198	0.82%

Interprétation : *La série après l'uniformisation n'est plus compressible par les deux premiers algorithmes utilisés. En revanche, l'algorithme par contexte, PAQ8o8, a pu engendrer un taux de compression positif.*

(0.82%), nous nous interrogeons naturellement sur la robustesse d'un tel résultat. Est-il possible que les variations du DJ aient été compressées par pure chance ? Autrement dit, une suite d'entiers indépendamment engendrée par un processus uniforme $U(0, 255)$ ²⁷, serait-elle compressible par l'algorithme PAQ8o8 ?

Pour résoudre cette question, nous avons simulé 100 suites de nombres entiers uniformément distribuées entre 0 et 255. Chaque série simulée a exactement la même longueur que la chronique du DJ (27432). Les algorithmes de compression sont appliqués sur chacune de ces chronologies aléatoires, mais aucun résultat positif n'a été constaté dans ce test de robustesse.

Sans pouvoir entièrement exclure la possibilité d'une compression due au hasard, le test de robustesse par simulation indique tout de même à quel point cela est difficile d'obtenir un taux de compression positif sur une série indépendamment et uniformément distribuée.

On peut donc conclure sans prendre beaucoup de risques que les algorithmes de compression peuvent détecter le phénomène des bouffées de volatilité. Ce résultat confirme encore une fois la relation "régularité-compressibilité" que nous avons défendue depuis le début du chapitre.

27. Ici, $U(0, 255)$ désigne la loi de distribution uniforme avec une valeur minimale de 0 et une valeur maximale de 255.

Or, la détection de régularités bien établies par des tests statistiques (ex. les bouffées de volatilité) n'est évidemment pas l'objectif ultime de l'approche algorithmique. Il serait intéressant d'avancer un pas de plus dans notre processus itératif d'effacement de structures pour vérifier si, derrière les faits stylisés bien connus, il y existe d'autres régularités détectables par des algorithmes de compression.

3.6.2 Degré d'aléa sans faits stylisés

Pour effacer les bouffées de volatilité, nous proposons ici un processus d'uniformisation progressive, qui uniformise les rentabilités du DJ en considérant leurs volatilités temporelles.

Le principe d'un tel processus est simple : en codant chaque rentabilité selon les variations à sa proximité temporelle, l'intervalle des rentabilités représenté par chaque nombre entier s'adaptera à la volatilité locale. Par exemple, pendant une période fortement volatile, l'entier 155 peut représenter une hausse de 5%, alors que pendant une période relativement "calme", ce même entier peut être associé à une hausse de 2%.

Ainsi, les valeurs extrêmes observées pendant les bouffées de volatilité (ex. la zone 9000 – 10000 dans la figure 3.18 page 152) sont "ramenées" vers zéro, alors que les valeurs observées pendant les périodes calmes (ex. la zone 13000 – 20000) seront "étirées" vers les extrêmes.

D'un point de vue technique, l'uniformisation progressive de la série du DJ, notée " $r_1, r_2, \dots, r_{27432}$ ", est réalisée en 3 étapes :

- Au départ, une fenêtre glissante de taille 512 est placée au début de la suite " $r_1, r_2, \dots, r_{27432}$ ". Les observations dans la fenêtre (cf. les premières 512 rentabilités de la série $r_1, r_2, \dots, r_{27432}$) sont calibrées dans 256 groupes de 2 en fonction de leur valeur en nombre réel. Les 256 groupes obtenus sont classés par ordre croissant, et à chaque rentabilité, on associe l'ordre du groupe la contenant. Ces premiers 512 entiers seront les premiers termes de la suite uniformisée, notée s .
- En décalant la fenêtre glissante d'une observation, r_1 sort de la fenêtre et r_{513} y entre. Toujours en fonction des rentabilités à l'intérieur de la fenêtre, nous uniformisons les rentabilités r_2, \dots, r_{513} . L'entier associé à r_{513} sera le 513ième terme de s .
- En répétant l'étape précédente jusqu'à la fin de la série " $r_1, r_2, \dots, r_{27432}$ ", s aura la même longueur que la suite des rentabilités. Surtout, s représente une volatilité relativement stable dans le temps. Une comparaison entre les

3.6. Une première application sur les données réelles : le cas de l'indice Dow Jones

figures 3.18 et 3.19 illustre l'effet de l'uniformisation progressive.

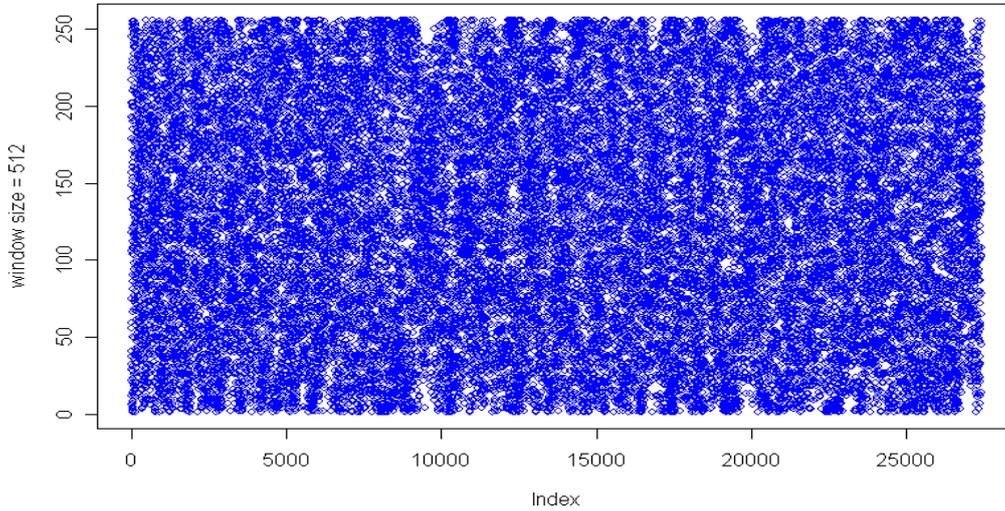


Fig. 3.19: Log-rentabilités du DJ après l'uniformisation progressive.

Les tests de compression sont conduits sur s , avec les résultats notés dans le tableau 3.22. On constate que même après effacement des bouffées de volatilité, la chronique s demeure compressible, certes peu, par PAQ8o8. La compressibilité de

Tab. 3.22: Tests de compression : Dow Jones après discrétisation progressive

Algorithme	taille du fichier	taux de compression
	27432	100%
Gzip	27498	-0,10%
Huffman	27756	-0,12%
PAQ8o8	27467	0,47%

Interprétation : *La série après uniformisation progressive est toujours compressible par PAQ8o8. Toutefois, le meilleur taux de compression réalisé est passé de 0,82% à 0,47%.*

s par PAQ8o8 implique l'existence de régularité dans les variations du DJ. Selon la théorie de la calculabilité, le taux de compression constaté (0,47%) mesure la quantité d'aléa contenue dans la suite des rentabilités. Ici, sa faible valeur correspond à un résultat classique en finance : même si la série du DJ ne peut pas être qualifiée comme relevant du "pur hasard", celle-ci dispose tout de même d'un degré d'aléa considérable.

Les variations du DJ étant fort imprévisibles, battre la "buy and hold" qui nécessite une prévision fiable des rentabilités négatives, sera donc une tâche ex-

trêmement difficile à accomplir. Ce résultat semble confirmer notre analyse des principaux courants théoriques en finance : quelle que soit la (ir)rationalité mise en avant, la “*buy and hold*” semble toujours imbattable à long terme. Nous attribuons ce résultat à la loi de l’offre et de la demande, qui nous apparaît valable même si les marchés boursiers représentent des phases d’optimisme ou de pessimisme.

Le taux de compression maximal, indicateur d’aléa (ou de régularité) calculé par la méthode algorithmique, mesure la prévisibilité d’une série donnée. Cette mesure offre non seulement un nouvel outil d’analyse dans les études des régularités financières, mais elle permet aussi de comparer quantitativement les différentes places boursières selon la structure de leur indice principal. Suivant cette logique, des études de comparaison seront menées dans les sections suivantes.

3.7 Application de la méthode aux marchés internationaux

Dans cette section, la méthode algorithmique est appliquée aux séries financières provenant de 3 marchés financiers différents : NYSE, NASDAQ et Euronext Paris. Au total, 111 titres sont choisis selon leur capitalisation boursière. Chaque titre choisi est caractérisé par 4 variables : le meilleur taux de compression après uniformisation totale, le meilleur taux de compression après uniformisation progressive²⁸, le nombre de rentabilités nulles, et la longueur de la série observée. En comparant les titres selon ces 4 critères, nous essaierons d’identifier les principaux facteurs explicatifs de leur degré de compressibilité.

L’étude présentée se déroule en 4 étapes : après la description des données utilisées, une analyse en composantes principales (ACP) est effectuée pour identifier les principaux facteurs qui caractérisent les 111 titres observés. Ensuite, les tests statistiques sont conduits pour vérifier la robustesse de l’ACP. Les résultats obtenus et leur signification économique sont discutés dans la dernière partie.

3.7.1 Présentation des données

Les données utilisées ici sont sans exception de fréquence “*tick-by-tick*”. C’est-à-dire que pour chaque titre observé, le nombre de prix enregistrés correspond exactement au nombre de transactions survenues pendant la période d’observation.

28. Comme nous l’avons présenté dans la section précédente, le processus d’uniformisation progressive est utilisé pour enlever les bouffées de volatilité dans les séries financières.

Pour représenter le NYSE, on étudie les composantes de l'indice SP100 satisfaisant les deux conditions suivantes : 1) être cotées sur NYSE, 2) présenter plus de 10000 observations pendant la période allant du 05/02/2001 au 09/02/2001. La deuxième condition garantit une longueur minimale pour les séries étudiées, afin que les algorithmes de compression puissent exploiter efficacement toutes les régularités dans leur champ d'application. La période d'observation est un choix arbitraire restreint par la disponibilité des données.

En appliquant ces critères, l'échantillon du NYSE contient 21 titres dont les noms sont exposés dans la colonne "Titre" du tableau M.3 de l'annexe M. La longueur de chaque série de rentabilités est reportée dans la colonne "Taille" du même tableau.

Pour le NASDAQ, ne sont étudiées que les composantes du NASDAQ-100 qui vérifient un nombre de transactions supérieur à 10000 pendant la même période d'observation que celle du NYSE. Cette contrainte sur le nombre d'observations réduit la taille de l'échantillon à 30 titres. Le nom et la longueur de chaque série retenue sont présentés dans le tableau M.4 de l'annexe M de la même manière que pour l'échantillon précédent. Les observations pour les deux marchés américains, l'un dirigé par les prix, l'autre par les ordres, sont extraites de la base de données "Trade and Quote" (TAQ).

En vue d'une comparaison franco-américaine, la place boursière Euronext est représentée par les composantes du CAC40 qui vérifient plus de 7000 transactions pendant la même période d'observation. Ici, la longueur maximale des séries retenues est moins élevée que pour les échantillons précédents²⁹, car les actions françaises engendrent, en moyenne, moins de transactions par jour que celles cotées aux États-Unis. La condition de sélection est légèrement modifiée pour obtenir un échantillon français de taille significative. Les titres sélectionnés sont présentés dans le tableau M.1 de l'annexe M, ainsi que la longueur de chaque série observée.

Les trois échantillons présentés ci-dessus partagent exactement la même période d'observation. Afin de pouvoir capter une éventuelle évolution temporelle des marchés financiers, un quatrième échantillon a été construit avec les observations plus récentes : sont utilisées à cet effet les composantes du CAC40 qui représentent au moins 10000 transactions du 24/11/2008 au 28/11/2008. Comme pour les trois échantillons précédents, cette nouvelle période d'observation correspond également à 5 jours de transactions consécutives. Ce dernier échantillon qui contient 38 actions a été construit, à l'aide des données fournies par la société NYSE-Euronext. Les titres retenus et la longueur de leur série "*tick-by-tick*" sont

29. On applique un seuil de 10000 observations pour les marchés américains, contre 7000 pour le marché français.

présentés dans le tableau M.2 de l'annexe M.

Remarquons que le nombre d'actions pour Euronext Paris est bien plus important pendant la période de 2008 qu'en 2001. Après 8 ans de développement, l'échange sur Euronext est nettement plus actif qu'en 2001. Le nombre de transactions en moyenne par jour et par titre pour les 40 composantes du CAC40 durant les deux périodes, est passé de 2347 à 8384³⁰. En conséquence, même en exigeant une longueur plus élevée, l'échantillon d'Euronext en 2008 contient plus de titres que celui en 2001.

Un dernier point à préciser concerne les périodes de transaction "inactives". Les ouvertures et fermetures des marchés boursiers sont souvent accompagnées d'une période où les ordres sont postés sans donner lieu à des fixations de prix. Durant cette période, les investisseurs enchérissent en émettant des ordres d'achat ou de vente, ces ordres émis ne seront exécutés qu'à la fin de cette période à un prix maximisant le nombre des ordres satisfaits (i.e. le cours d'ouverture ou de fermeture). Comme tous les ordres émis durant ces enchères sont exécutés au même prix, on observe souvent une longue série de rentabilités zéros après chaque ouverture et avant chaque fermeture du marché.

Pour éviter l'influence de ces périodes sur le degré d'aléa des suites de rentabilités, les 5 premières minutes de chaque journée de transactions ne sont pas prises en compte dans nos analyses. Autrement dit, dans notre base de données, les journées de transactions débutent à 9h35 et finissent à 16h30 pour le NYSE et le NASDAQ. Pour Euronext, chaque journée de transactions commence à 9h05 et finit à 17h30.

3.7.2 Tests de compression et résultats obtenus

Pour quantifier le degré d'aléa de chaque titre observé, les tests de compression ont été conduits sur chaque série de rentabilités successives. Ces tests de compression se déroulent en 4 étapes :

- Dans un premier temps, chaque série est soumise au processus d'uniformisation présenté dans la section précédente. À l'issue de cette première étape, on obtient, pour l'année 2001, un échantillon de
 - 21 suites de nombres entiers pour le NYSE,
 - 30 pour le NASDAQ,
 - 22 pour Euronext.

Un quatrième échantillon pour Euronext de 38 séries de rentabilités correspondant aux observations de 2008.

30. Moyennes obtenues avec les données "tick-by-tick" fournies par Euronext-NYSE.

- Ensuite, les suites de nombres entiers sont compressées par les 3 algorithmes présentés dans la section précédente³¹. Pour chacune de ces suites, le taux de compression maximal est noté dans les colonnes “Taux1” des tableaux M.1, M.2, M.3 et M.4 de l’annexe M.

Il est à noter dès à présent que tous les taux de compression maximaux ont été réalisés avec l’algorithme par contexte PAQ8o8, présenté dans la section 3.4.2 page 124.

- Comme dans le cas des rentabilités journalières, les variations “*tick-by-tick*” manifestent également un phénomène de bouffées de volatilité. Il est donc nécessaire d’effacer ce phénomène en avançant un pas de plus dans le processus d’effacement de régularités, comme on le fait pour le DJ dans la section précédente.

À cet effet, la méthode de fenêtre glissante présentée dans la section 3.6.2 (page 154) est appliquée aux séries des rentabilités réelles.

- Après l’effacement des bouffées de volatilité, les suites de rentabilités sont à nouveau compressées par les 3 algorithmes de compression cités ci-dessus. Les meilleurs taux de compression obtenus, sont exposés dans les colonnes “Taux2” des tableaux M.1, M.2, M.3 et M.4.

3.7.3 Traitement des rentabilités zéros

Comparées avec les données journalières, une particularité des rentabilités “*tick-by-tick*” consiste en une présence excessive de valeurs nulles. Sur une échelle de temps de quelques secondes, les transactions successives d’un titre donné partagent souvent le même prix de transaction. Comme chaque échange est enregistré dans les bases de données “*tick-by-tick*”, les rentabilités successives calculées à cette fréquence représentent souvent une proportion très élevée de valeurs nulles. Par exemple, le ratio $\frac{\text{rentabilités nulles}}{\text{nombre d'observations total}}$ pour les 38 composantes de CAC40 durant la période d’observation en 2008, varie entre 45,5% et 71,4% avec une moyenne de 56,3% et un écart-type de 4,75%.

Cette présence excessive de rentabilités nulles est une difficulté technique à résoudre dans le processus d’uniformisation. Comme développé dans la section 3.5.2 (page 135), pour calculer l’entier correspondant à chaque rentabilité historique, les variations de prix doivent être classées par ordre croissant. Or, si chaque série contient environ 50% de valeurs nulles, la moitié des 256 nombres entiers utilisés dans l’uniformisation correspondra aux rentabilités nulles. Comment associer ces différents nombres entiers aux rentabilités nulles qui portent toutes le même rang? Le traitement de ces rentabilités nulles risque de modifier le degré d’aléa des suites uniformisées, et de fausser les conclusions de cette étude. Pour

31. Pour rappeler, Gzip, Huffman et PAQ8o8.

éviter toute régularité rajoutée par le processus d'uniformisation, l'association des rentabilités nulles aux nombres entiers s'est effectuée de manière aléatoire.

Les paragraphes suivants détaillent et discutent de cette procédure :

Soit A , une série de 10000 observations qui contient 5000 rentabilités nulles, 2500 rentabilités négatives et 2500 positives. Si on souhaite coder ces rentabilités avec 256 entiers, l'uniformisation de A nécessite une correspondance entre les rentabilités nulles et tous les entiers se trouvant dans l'intervalle $[64, 171]$. Pour éviter d'introduire de nouvelles régularités, chaque rentabilité nulle est associée à l'un des entiers entre $[64, 171]$ de manière aléatoire. Autrement dit, une rentabilité nulle a autant de chance d'être associée à chacun des entiers dans l'intervalle, et ce, quel que soit son ordre d'apparition dans A .

Pour une série de rentabilités aléatoires, le traitement des rentabilités nulles proposé ci-dessus ne doit pas avoir d'influence significative sur la compressibilité de la série uniformisée. En revanche, pour une série de rentabilités où l'apparition des zéros aide à la prédiction des termes futurs, ce traitement rendra plus difficile à repérer ces structures liées aux rentabilités nulles, cela augmentera donc le degré d'aléa de la suite uniformisée.

Par exemple, soit une série de rentabilités s , qui vérifie la régularité suivante : chaque séquence de 4 rentabilités nulles est suivie par une augmentation. Avant uniformisation, toutes les séquences de 4 rentabilités nulles sont considérées comme un seul et unique contexte de prédiction. La régularité supposée est facilement identifiée par un algorithme de compression, qui réduira la taille de s en l'exploitant. En revanche, après uniformisation, chaque rentabilité nulle est remplacée par l'un des entiers entre $[64, 171]$ au hasard. Toutes les séquences de 4 rentabilités nulles ne seront plus considérées comme un seul contexte de prédiction, et la régularité dans s sera difficile à repérer. Le traitement des rentabilités zéros a donc augmenté le degré d'aléa de s .

Pour une série de rentabilités comme s , l'impact du traitement dépend de la proportion des rentabilités nulles sur l'ensemble des variations de prix, son taux de compression maximal doit être négativement corrélé avec cette dernière proportion.

Pour tester cette hypothèse, la proportion de rentabilités nulles dans chaque série de rentabilités est calculée et exposée dans les colonnes “%zéro” des tableaux M.1, M.2, M.3 et M.4. La relation entre ces proportions et les degrés d'aléa observés sera testée dans la section “Signification statistique des résultats obtenus” (page 164).

3.7.4 Méthodes et analyses

Après les tests de compression, chaque titre financier est caractérisé par 4 variables différentes : “Taille”, “Taux1”, “Taux2”, et “%zéro”. En fusionnant les 4 marchés analysés, on obtient une base de données globale de 111 actions.

Existe-il des relations de dépendance entre ces variables ? Quels sont les principaux facteurs qui influencent le degré d’aléa d’une série de rentabilités ? Les actions cotées sur le même marché, disposent-elles de taux de compression similaires ? Telles sont les questions qu’on espère résoudre en explorant la base de données obtenue. Le but de cette section est d’étudier ces questions à l’aide de différents tests statistiques, et de discuter des résultats obtenus en proposant des explications économiques.

Après une ACP dont les détails sont exposés dans l’annexe N, on extrait deux composantes principales qui expliquent jusqu’à 85,79% de la variance totale de l’échantillon. Les composantes 3 et 4 qui expliquent, en totalité, moins de 15% de la variance totale, seront donc négligées dans les futures analyses.

Ensuite, pour découvrir le sens économique des composantes principales, les “scores” de chaque titre sur ces deux composantes sont respectivement calculés selon la méthode détaillée dans l’annexe N.

Les résultats obtenus sont présentés dans la figure 3.20 page 162, l’axe X représentant les “scores” des actions sur la composante 1 et l’axe Y les “scores” sur la composante 2. L’image illustre le positionnement de chaque action en considérant uniquement les deux composantes principales.

On constate dans la figure 3.20, trois groupes d’actions distincts : les actions numérotées de 61 à 81, provenant du NYSE, sont concentrées dans la partie en bas à gauche de la figure. Les titres numérotés entre 1 et 60, qui appartiennent tous à Euronext, se trouvent exclusivement sur le côté droit de l’image. Enfin, les numéros supérieurs à 82, attribués aux actions cotées sur le NASDAQ, occupent principalement la partie en haut à gauche de la figure 3.20.

Cette segmentation est confirmée par les figures 3.21(a) et 3.21(b), qui présentent les “scores” des actions sur chaque composante principale. La figure 3.21(a) montre l’effet de la composante 1 sur la segmentation : elle permet de distinguer les actions cotées en France (i.e., numéro 1-60) de celles qui le sont aux États-Unis (numéro 61-111). Alors que la figure 3.21(b) différencie les titres cotés sur le NYSE (numéro 61-81) de ceux des autres marchés.

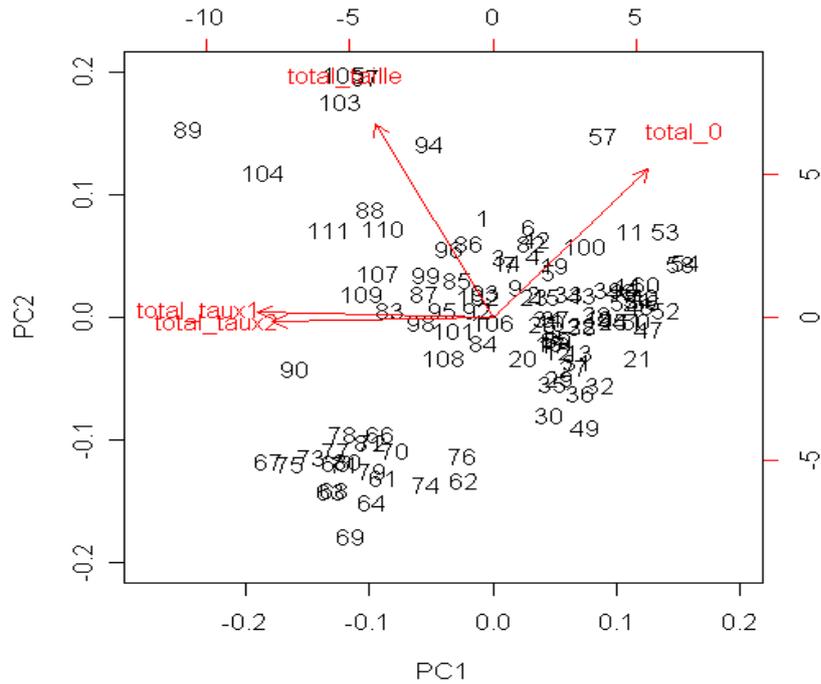
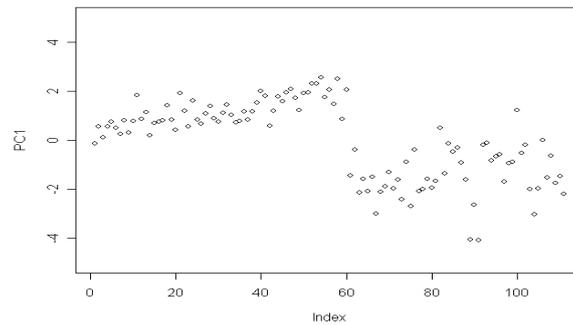


Fig. 3.20: Projection des titres sur les deux composantes principales.

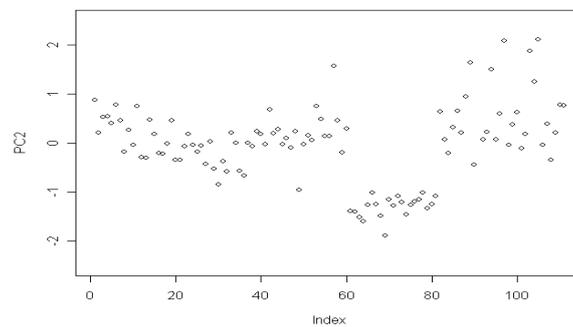
Dans cette figure, les actions cotées sur Euronext sont représentées par les numéros 1 à 22 pour la période 2001, et 23-60 pour la période 2008, celles cotées sur le NYSE correspondent aux numéros entre 61 et 81. Les actions cotées sur le NASDAQ sont présentées par les numéros de 82 à 111. Les flèches sont les projections des 4 variables d'origine dans l'espace vectoriel défini par les composantes 1 et 2.

Soit θ , l'angle formé par un axe et une flèche, alors, la corrélation entre la composante principale représentée par l'axe et la variable désignée par la flèche est calculée par $\cos \theta$. Par exemple, l'axe X et la flèche "Total taux1" forment un angle d'environ 175° , alors, la corrélation entre la composante 1 et la variable *taux1* est de $-1 < \cos 175^\circ < -0,9$. En fonction de l'angle formé par composante 1 et les 4 flèches en rouge, on déduit que la variable "total_0" (i.e. %zéros) est positivement corrélée avec la composante, alors que les autres variables vérifient une corrélation négative avec celle-ci. Les angles formés entre les 4 flèches et l'axe Y impliquent que les variables "Taille" et "%zéro" sont positivement corrélées avec la composante 2.

3.7. Application de la méthode aux marchés internationaux



(a) “Scores” des titres selon la composante 1



(b) “Scores” des titres selon la composante 2

Fig. 3.21: “Scores” des actions sur les composantes 1 et 2.

Dans ces figures, chaque point représente un titre étudié. L’axe X donne le numéro identifiant chaque action, l’axe Y montre son “score” pour la composante concernée. Les actions numérotées de 1 à 22 sont cotées sur Euronext et observées pendant la période en 2001, de 23 à 60 sont les composantes de CAC40 observées en 2008, de 61 à 81 les représentants de NYSE, et de 82 à 111 ceux de NASDAQ.

Désormais, les significations économiques des deux composantes principales deviennent plus claires : PC1 distingue Euronext des marchés américains, alors que PC2 marque la différence entre NYSE, le seul marché dirigé par les prix, et les autres marchés observés.

Fondées sur ces deux composantes principales, 5 propositions sont formulées pour interpréter Base I.

1. Le degré d'aléa de chaque titre est mesuré par deux variables différentes : "Taux1" qui désigne le taux de compression maximal en présence des bouffées de volatilité et "Taux2", qui donne ce même taux en absence de ce fait stylisé (une fois les "bouffées de volatilité" effacées). Dans la figure 3.20, on constate une forte corrélation entre ces deux mesures d'aléa, ce qui indique une certaine cohérence entre les deux mesures de complexité.

Une série de rentabilités plus complexe en présence des bouffées de volatilité, a tendance à le rester après l'uniformisation progressive. Cela est plutôt un argument en faveur de notre méthode de transformation.

2. Les titres cotés aux États-Unis ont des degrés d'aléa moins élevés que ceux échangés en France.
3. La façon dont un marché est organisé (dirigé par les prix ou par les ordres) n'a pas d'impact sur la compressibilité des rentabilités, car "Taux1" et "Taux2" sont très peu corrélés avec PC2.
4. Le taux de rentabilités nulles et la longueur de la série de rentabilités sont influencés par les deux composantes principales.
5. La corrélation entre le taux de rentabilités nulles (ou la taille) et le degré d'aléa d'une série donnée, est principalement expliquée par le pays de cotation.

La méthode d'ACP s'appuie sur les corrélations observées. Comme Base I n'est qu'un échantillon tiré d'une population, les corrélations observées nécessitent des contrôles de signification. C'est pourquoi ces conclusions tirées de l'ACP sont soumises aux tests statistiques développés ci-dessous.

Signification statistique des résultats obtenus

Différents tests sont réalisés pour contrôler la signification statistique des 5 propositions évoquées ci-dessus. Ces tests statistiques, regroupés dans les 5 paragraphes suivants, confirment presque toutes les propositions obtenues, sauf l'effet du pays de cotation sur la taille des séries observées.

1. La corrélation entre les variables *Taux1* et *Taux2* est significative.

Une régression linéaire est effectuée entre les 111 observations des deux variables concernées. La signification statistique du coefficient devant la variable indépendante est interprétée comme celle de leur corrélation linéaire.

En consultant le tableau de résultat (i.e. tableau 3.23), on peut rejeter l'hypothèse nulle "le paramètre devant la variable *Taux2* est nul" avec un risque inférieur à 0,1%. Ce résultat confirme donc la première conclusion tirée de l'ACP : les ac-

Tab. 3.23: Régression linéaire entre *Taux1* et *Taux2*

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.0000	0.0425	0.00	1.0000
<i>Taux2</i>	0.8954	0.0426	21.00	$< 2e - 16^{***}$

L'équation linéaire obtenue est $taux1 = 0,8954 \cdot taux2$, dont le coefficient 0,8954 est significativement différent de 0 au seuil de 0,1%.

"*" correspond au seuil de signification à 5%, "***" à 1%, et "****" à 0,1%.

tions les plus compressibles en présence des bouffées de volatilité le resteront après l'effacement du fait stylisé.

Grâce à cette forte corrélation, l'usage de *Taux2* comme ultime indicateur du degré d'aléa ne doit pas modifier radicalement les conclusions finales, notamment quand il s'agit de comparer les différents marchés financiers.

2. Il existe une différence significative entre la compressibilité des titres cotés en France et aux États-Unis.

Dans la Base I, il existe 111 titres observés, dont 60 sont cotés en France et 51 aux États-Unis. Pour vérifier la différence entre ces deux zones géographiques, les *Taux2* obtenus pour tous les titres sont séparés en deux parties selon leur pays de cotation. Les deux échantillons obtenus, nommés respectivement *Taux2_France* et *Taux2_Usa*, sont comparés à l'aide d'un test de Kolmogorov-Smirnov (désormais "K-S").

Dans le tableau 3.24 on constate une différence significative à 0,01% entre *Taux2_France* et *Taux2_Usa*. Ce qui confirme la deuxième proposition émise à l'issue de l'étude ACP.

3.7. Application de la méthode aux marchés internationaux

Tab. 3.24: K-S test effectué sur Taux2_France et Taux2_Usa

Two-sample Kolmogorov-Smirnov test	
Data :	Taux2_France and Taux2_Usa
D = 1	P-value < $2.2e - 16^{***}$

Hypothèse nulle : les deux échantillons proviennent de la même population.
 “*” correspond au seuil de signification à 5%, “***” à 1%, et “****” à 0,1%.

3. La façon dont les transactions sont organisées n’affecte pas la compressibilité des titres cotés dans un même pays.

Deux étapes sont envisagées pour tester l’homogénéité au sein de l’échantillon Taux2_France et Taux2_Usa :

- Dans un premier temps, le test de Shapiro est utilisé pour vérifier la normalité de la distribution des deux échantillons. Les résultats exposés respectivement dans les tableaux 3.25 et 3.26 montrent un risque trop élevé pour rejeter l’hypothèse nulle associée à la normalité de la distribution. Les taux de compression observés au sein d’un même pays sont normalement distribués, cela confirme, dans certaines mesures, l’homogénéité des deux échantillons Taux2_France et Taux2_Usa. De plus, comme Taux2_Usa contient un marché dirigé par les prix (NYSE) et un marché dirigé par les ordres (NASDAQ), on fournit un premier argument soutenant l’indépendance entre le principe d’organisation et la compressibilité des titres.

Tab. 3.25: Test de Shapiro conduit sur Taux2_France

Shapiro-Wilk normality test	
Data :	Taux2_france
W = 0.9624	P-value = 0.06166

Hypothèse nulle : l’échantillon est distribué selon une loi normale.

Tab. 3.26: Test de Shapiro conduit sur Taux2_Usa

Shapiro-Wilk normality test	
Data :	Taux2_usa
W = 0.9862	P-value = 0.8113

Hypothèse nulle : l’échantillon est distribué selon une loi normale.

- Ensuite, Taux2_France est divisé en deux groupes ; le groupe des taux2 observés en 2001, nommé *France2001*, et le groupe des taux2 observés en 2008, nommé *France2008*. Taux2_Usa est divisé en deux sous-échantillons selon le marché d’origine des taux2 : le sous-échantillon *NYSE* qui contient les taux2 observés sur le marché NYSE, et le sous-échantillon *NASDAQ* qui contient des degrés d’aléa observés sur le marché NASDAQ.

3.7. Application de la méthode aux marchés internationaux

Les tests de K-S sont conduits pour comparer *France2001* et *France2008*, ainsi que *NYSE* et *NASDAQ*. Les résultats obtenus, présentés dans les tableaux 3.27 et 3.28, indiquent l'absence de différence significative entre *France2001* et *France2008*, ainsi qu'entre *NYSE* et *NASDAQ*.

Tab. 3.27: K-S test effectué sur *France2001* et *France2008*
Two-sample Kolmogorov-Smirnov test

Data :	<i>France2001</i> and <i>France2008</i>
D = 0.1603	P-value = 0.8001

Hypothèse nulle : les deux échantillons proviennent de la même population.

Tab. 3.28: K-S test effectué sur *NYSE* et *Nasdaq*
Two-sample Kolmogorov-Smirnov test

Data :	<i>NYSE</i> and <i>Nasdaq</i>
D = 0.2619	P-value = 0.3650

Hypothèse nulle : les deux échantillons proviennent de la même population.

La stabilité des Taux2 observés dans un même pays rend légitime notre tentative de comparer les marchés financiers au regard de cet indicateur d'aléa. L'absence de différence significative entre *NYSE* et *NASDAQ* confirme en même temps la troisième proposition à l'issue de l'ACP : "La façon dont un marché est organisé (dirigé par les prix ou par les ordres) n'a pas d'influence sur la compressibilité des variations de prix." C'est un résultat qui, en microstructure, est intéressant puisqu'il suggère que le choix d'organisation de la place n'influence pas son "efficacité" de façon évidente.

4. "%zéro" et "Taille" sont influencées en même temps par les deux composantes principales.

- **L'effet du pays de cotation sur "%zéro"** : Les 111 observations de la variable "%zéro" sont séparées en 2 parties selon leur marché d'origine. "Usa_taux0" dénote l'échantillon contenant les taux de rentabilités nulles observés aux États-Unis, et "France_taux0", l'échantillon des taux de rentabilités nulles observés en France. Le test de K-S est réalisé pour comparer statistiquement "Usa_taux0" et "France_taux0". Le tableau 3.29 reporte une différence significative à moins de 0,1% entre les deux sous-échantillons. L'influence du pays de cotation sur le taux de rentabilités nulles est ainsi confirmée.

- **L'effet du principe d'organisation sur "%zéro"** : Comme le NYSE est le seul marché dirigé par les prix dans la Base I, la relation entre les taux des rentabilités nulles et l'organisation des transactions dépend de la différence entre

3.7. Application de la méthode aux marchés internationaux

Tab. 3.29: K-S test effectué sur “Usa_taux0” et “France_taux0”

Two-sample Kolmogorov-Smirnov test	
Data :	“Usa_taux0” and “France_taux0”
D = 0.4569	P-value = $9.688e - 06^{***}$

Hypothèse nulle : les deux échantillons proviennent de la même population.
“*” correspond au seuil de signification à 5%, “**” à 1%, et “***” à 0,1%.

l'échantillon “Quote_taux0” qui contient les “%zéro” des titres cotés sur le NYSE, et l'échantillon “Ordre_taux0” qui regroupe les autres observations de “%zéro”.

Le test de K-S est utilisé pour tester la signification de cette différence. Le résultat, exposé dans le tableau 3.30, révèle une différence significative à moins de 0,1% entre les deux échantillons analysés. Ce qui confirme l'influence de la méthode organisationnelle sur le taux de rentabilités nulles.

Tab. 3.30: K-S test effectué sur “Quote_taux0” et “Ordre_taux0”

Two-sample Kolmogorov-Smirnov test	
Data :	“Quote_taux0” and “Ordre_taux0”
D = 0.9667	P-value < $2.2e - 16^{***}$

Hypothèse nulle : les deux échantillons proviennent de la même population.

- **L'effet du pays de cotation sur l'indicateur “Taille”** : Comme pour les taux de rentabilités zéros, la relation entre la taille des séries observées et la manière dont les transactions sont organisées est soumise aux mêmes tests de signification. Dans le tableau 3.31 on présente la comparaison entre “Usa_taille” et “France_taille”, deux échantillons qui contiennent exactement les mêmes titres que “Usa_taux0” et “France_taux0”.

En revanche, le résultat de cette comparaison s'est révélé non significatif au seuil de 5%. Ceci suggère une absence d'impact du pays de cotation sur la taille des séries observées.

Tab. 3.31: K-S test effectué sur “Usa_taille” et “France_taille”

Two-sample Kolmogorov-Smirnov test	
Data :	“Usa_taille” and “France_taille”
D = 0.2108	P-value = 0.1725

Hypothèse nulle : les deux échantillons proviennent de la même population.

- **L'effet du principe d'organisation sur "Taille"** : L'impact du principe d'organisation sur la taille des séries observées est testé en comparant les groupes "Quote_taille" et "Ordre_taille0". Ici, chacun de ces deux échantillons contient exactement les mêmes titres que "Quote_taux0" et "Ordre_taux0".

Le résultat de cette comparaison, présenté dans le tableau 3.32, confirme la différence significative entre "Quote_taille" et "Ordre_taille". Autrement dit, la façon dont les transactions sont organisées sur un marché influence significativement la longueur des séries observées.

Tab. 3.32: K-S test effectué sur "Quote_taille" et "Ordre_taille"

Two-sample Kolmogorov-Smirnov test	
Data : Quote_taille" and "Ordre_taille"	
D = 0.527	P-value = 0.0001562***

Hypothèse nulle : les deux échantillons proviennent de la même population.
 "*" correspond au seuil de signification à 5%, "**" à 1%, et "***" à 0,1%.

L'impact du principe d'organisation sur "Taille" est probablement due aux efforts des "market makers" qui regroupent les "petits" ordres similaires. En effet, dans un marché dirigé par les ordres, le carnet d'ordres enregistre tous les ordres passés concernant un titre donné. S'il existe 10 ordres qui achètent la même action au même prix, le carnet d'ordres enregistre 10 ordres différents sans aucun regroupement. Et ce, quelle que soit la quantité réclamée par chaque ordre. Quand un vendeur accepte le prix proposé, les 10 ordres seront exécutés respectivement dans leur ordre d'apparition, en laissant dans la série de prix "tick-by-tick", 10 transactions successives au même prix.

En revanche, sur un marché dirigé par les prix, les ordres similaires sont regroupés par les "market makers" pour simplifier la rencontre des ordres. Un "market maker" recevant 10 ordres voulant acheter la même action au même prix a tendance à les regrouper pour n'effectuer qu'une seule transaction. Ainsi, l'exécution de ces 10 ordres n'engendrera qu'une seule transaction au prix proposé. A cause d'un tel regroupement d'ordres, les séries de rentabilités "tick-by-tick" observées pendant la même période sont souvent plus longues sur un marché dirigé par les ordres que par les prix. Ce raisonnement pourrait aussi expliquer la relation entre "%zéro" et l'organisation des marchés.

5. La corrélation entre “Taille” (“%zéros”) et “Taux2” est robuste, mais la relation “%zéros-Taux2” est expliquée par le pays de cotation

Pour vérifier la robustesse de ces deux corrélations, des régressions linéaires sont effectuées entre les deux couples de variables. La signification du coefficient devant la variable indépendante est interprétée comme la robustesse de la corrélation testée.

Les deux régressions linéaires sont exposées respectivement dans les tableaux 3.33 et 3.34. Le t-test effectué sur le coefficient devant Taille2 confirme la signification des corrélations analysées.

Tab. 3.33: Régression linéaire effectuée entre %zéros et Taux2

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.0000	0.0864	-0.00	1.0000
Taux2	-0.4240	0.0867	-4.89	3.53e - 06***

Le coefficient devant *Taux2* est significatif à moins de 0,1%.

“*” correspond au seuil de signification à 5%, “**” à 1%, et “***” à 0,1%.

Tab. 3.34: Régression linéaire effectuée entre Taille et Taux2

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.0000	0.0909	0.00	1.0000
Taux2	0.3006	0.0914	3.29	0.0013**

Le coefficient devant *Taux2* est significatif à moins de 1%.

“*” correspond au seuil de signification à 5%, “**” à 1%, et “***” à 0,1%.

Or, il est encore trop tôt de conclure sur la causalité entre la “Taille” (“%zéros”) d’une série donnée et son degré d’aléa *Taux2*. En effet, si le pays de cotation influence “Taille” (“%zéros”), et qu’il existe une relation étroite entre le pays de cotation et les “Taux2”, alors, la corrélation observée dans les deux tableaux ci-dessus peut être une simple conséquence de ces deux causalités déjà constatées.

Pour vérifier cette hypothèse, les tests de K-S sont conduits sur les observations d’un même pays. Si la relation entre Taux2 et Taille (%zéros) disparaît au sein d’un même pays, on considérera que les corrélations constatées dans les tableaux 3.33 et 3.34 sont expliquées par le pays de cotation.

À cet effet, les 111 observations de Taux2 sont d’abord réparties en 2 groupes selon leur pays de cotation, le groupe “France_taux2” qui contient les “Taux2” observés en France et “Usa_taux2” qui regroupe les observations réalisées aux

3.7. Application de la méthode aux marchés internationaux

États-Unis. Ensuite, “France_taux2” et “Usa_taux2” sont respectivement séparés en deux parties selon les tailles des séries correspondantes. Le sous-échantillon “France_taux2_long” (resp. “France_taux2_court”) contient les “Taux2” des séries les plus longues (resp. courtes) observées France. Les sous-échantillons “Usa_taux2_long” et “Use_taux2_court” sont construits selon ce même principe.

La différence entre “France_taux2_long” et “France_taux2_court”, et celle entre “Usa_taux2_long” et “Usa_taux2_court” sont vérifiées avec les tests de K-S dont les résultats sont respectivement exposés dans les tableaux 3.35 et 3.36.

Tab. 3.35: K-S test effectué entre “France_taux2_long” et “France_taux2_court”

Two-sample Kolmogorov-Smirnov test	
Data :	“France_taux2_long” and “France_taux2_court”
D = 0.4227	P-value = 0.00634**

Hypothèse nulle : les deux échantillons proviennent de la même population.
 “*” correspond au seuil de signification à 5%, “**” à 1%, et “***” à 0,1%.

Tab. 3.36: K-S test effectué entre “Usa_taux2_long” et “Use_taux2_court”

Two-sample Kolmogorov-Smirnov test	
Data :	“Usa_taux2_long” and “Use_taux2_court”
D = 0.3662	P-value = 0.06559*

Hypothèse nulle : les deux échantillons proviennent de la même population.
 Le P-value constaté ici est légèrement au dessus du seuil de 5%. Cependant, comme “0,06559” est très proche du seuil de 5%, on rejette l’hypothèse nulle.

On rapporte dans ces tableaux une différence significative entre “France_taux2_long” et “France_taux2_court” (“Usa_taux2_long” et “Use_taux2_court”). Cela confirme donc l’influence de la taille sur le degré d’aléa des séries observées.

Une explication probable du phénomène se trouve dans la relation théorique entre la taille d’un fichier donné et le taux de compression réalisé par un algorithme. D’une manière générale, plus un fichier de texte est long, plus le taux de compression réalisé s’approchera du taux de compression théorique. Tout fichier compressé par un algorithme sans perte contient un en-tête dont la taille dépend uniquement de l’algorithme utilisé. Quelle que soit la technique de compression utilisée, cet en-tête est toujours nécessaire car elle contient les informations indispensables à la décompression.

Soit p un algorithme de compression qui économise 1 bit d’information sur 8 d’un fichier de taille s . Le taux de compression **théorique** est $\rho = \frac{1}{8}$. Supposons que l’en-tête du fichier compressé soit de longueur e ; alors, le taux de compression

3.7. Application de la méthode aux marchés internationaux

réalisé par p est calculé par $t = 1 - \frac{e+s(1-\rho)}{s} = \rho - \frac{e}{s}$. Ce taux de compression réalisé augmente avec la taille du fichier s .

La relation entre “%zéros” et “Taux2” des titres cotés dans un même pays est vérifiée en suivant la même logique que dans les tests portant sur la relation “Taille-Taux2”. Cette fois-ci, les échantillons “France_taux2” et “Usa_taux2” sont respectivement divisés en deux sous-échantillons selon le taux des rentabilités nulles.

Les 4 sous-échantillons ainsi obtenus, nommés respectivement “France_taux2_élevé”, “France_taux2_modéré”, “Usa_taux2_élevé” et “Usa_taux2_modéré”, sont soumis aux tests de K-S dont les résultats sont exposés dans les tableaux 3.37 et 3.38.

Tab. 3.37: K-S test effectué entre “France_taux2_élevé” et “France_taux2_modéré”

Two-sample Kolmogorov-Smirnov test	
Data :	“France_taux2_élevé” and “France_taux2_modéré”
D = 0.2102	P-value = 0.4521

Hypothèse nulle : les deux échantillons proviennent de la même population.

Tab. 3.38: K-S test effectué entre “Usa_taux2_élevé” et “Usa_taux2_modéré”

Two-sample Kolmogorov-Smirnov test	
Data :	“Usa_taux2_élevé” and “Usa_taux2_modéré”
D = 0.2462	P-value = 0.4227

Hypothèse nulle : les deux échantillons proviennent de la même population.

On ne constate aucune différence significative entre les sous-échantillons provenant du même pays. Cela confirme notre dernière proposition à l’issue de l’ACP : au sein d’un même pays de cotation, le taux de rentabilités nulles n’influence pas la compressibilité des séries observées, la corrélation entre “% zéros” et “Taux2” observée dans la figure 3.20 (page 162) est expliquée par la première composante principale : le pays de cotation.

3.7.5 Résultats et discussions

Le taux de compression maximal d’une série de rentabilités indique son degré de structuration. En adoptant une telle logique, cette étude vise à mesurer le degré aléa de différentes séries financières, et en comparant les résultats obtenus, à identifier les principaux facteurs qui influencent cet indicateur de hasard.

À cet effet, 3 marchés sont observés à une fréquence *“tick-by-tick”* : le NYSE, le NASDAQ, et Euronext, durant la période allant du 05/02/2001 au 09/02/2001. Euronext est également étudié pendant la période du 24/11/2008 au 28/11/2008. Le taux de compression maximal est calculé pour chacun des 111 titres observés, respectivement avant et après effacement des bouffées de volatilité.

Après une série d’analyses et de comparaisons, 5 conclusions peuvent être tirées :

1. Bien que le taux de compression maximal connaisse une baisse considérable après effacement des bouffées de volatilité, une série plus compressible avant cette opération a une forte tendance à le rester après. Cela est sans doute un argument en faveur du processus d’uniformisation progressive qui permet de stabiliser la volatilité des séries de rentabilités. Ce résultat rend légitime l’usage de “Taux2” comme l’indicateur ultime du degré d’aléa dans les chroniques.
2. Les degrés d’aléa des titres cotés sur le même marché sont stochastiquement identiques. C’est justement cette stabilité qui justifie la comparaison des marchés financiers selon la compressibilité des séries de rentabilités.
3. La longueur d’une série donnée n’est pas influencée par son pays de cotation. Celle-ci est fort sensible à la manière dont les transactions sont organisées (marché dirigé par les ordres ou par les prix).
En revanche, la variable “%zéro” est aussi bien influencée par le pays de cotation que par la méthode d’organisation.
4. “Taille” et “%zéro” ont toutes un effet sur la compressibilité des séries étudiées. Or, si la dépendance entre la longueur d’une série donnée et sa prévisibilité est expliquée par la présence de l’en-tête dans des fichiers compressés, la relation entre “%zéro” et “Taux2” doit son explication au changement du pays de cotation.
5. La compressibilité d’un titre donné est significativement influencée par le pays de cotation. Plus précisément, les marchés américains disposent d’une “prévisibilité théorique” plus élevée que le marché français. Et ce, quelle que soit la période d’observation, et quelle que soit la manière dont les transactions sont organisées.

Le dernier point - constat d’un degré d’aléa plus élevé en France qu’aux États-Unis - mérite une discussion particulière.

Le degré d’aléa d’un titre donné indique la difficulté à prévoir ses rentabilités futures. D’une manière générale, plus une série est difficile à prévoir, plus la *“buy and hold”* sera difficile à battre. Or, “battre le marché” nécessite la prédiction des “baisses”, et compte tenu des algorithmes utilisés, la compression des séries financières ne va pas toujours de pair avec une telle prédiction.

Ainsi, pour bien interpréter les taux de compression obtenus, il serait idéal de connaître précisément les régularités exploitées par l'algorithme pour comprendre leur effet sur l'efficience informationnelle. C'est pourquoi nous proposons ici une discussion sur l'origine de la différence de compressibilité entre la France et les États-Unis.

Avant d'émettre des explications plausibles, on va d'abord citer les facteurs qui ne peuvent pas être à l'origine du problème.

1. La proportion des rentabilités nulles dans une série donnée n'influence pas sa prévisibilité. Selon l'hypothèse émise dans la section 3.7.3 page 159, ce résultat indique l'absence de modèles de prédiction fondés sur l'ordre d'apparition des rentabilités nulles.
2. Parmi les 3 marchés analysés, le NYSE est dirigé par les prix, alors que NASDAQ et Euronext sont tous deux dirigés par les ordres. La différence en prévisibilité serait-elle expliquée par la présence des "*market makers*" sur le NYSE ? En consultant le tableau 3.28 (page 167), on constate une différence non significative entre les compressibilités des titres cotés sur le NASDAQ et sur le NYSE. Si l'organisation des transactions était à l'origine du phénomène, une différence de prévisibilité entre NASDAQ et NYSE devrait être constatée. Cette hypothèse est donc aussi écartée.

Après avoir limité ces premières conjectures, la compression d'une série de rentabilités s après le processus d'uniformisation progressive peut se produire dans 2 cas de figure :

1. s est autocorrélée. La littérature en micro structure des marchés offre deux explications à l'auto-corrélation des rentabilités "*tick-by-tick*" :
 - l'existence de la fourchette de cotation (i.e. le "*bid ask spread*"). Selon le modèle de Roll (1984) (présenté dans la section 2.1.1 page 59), l'auto-corrélation entre les rentabilités "*tick-by-tick*" est autant plus élevé que la valeur du "*bid ask spread*" l'est elle même.
 - la discontinuité des rentabilités causée par la concentration des prix autour des valeurs entières ou demies (Glosten et Milgrom 1985). Par exemple, les prix comme 10, 9, 50 ou $\frac{21}{2}$, $\frac{19}{2}$, $\frac{101}{2}$ apparaissent beaucoup plus souvent dans le carnet d'ordres que 10, 1, 9, 2, 50, 7 etc. Selon Harris (1990), cette discontinuité engendre l'auto-corrélation des rentabilités "*tick-by-tick*".

Les places boursières américaines exigeraient-elles un "*bid ask spread*" plus important que le marché français ? À notre connaissance, aucune étude sur ce sujet n'est actuellement disponible. Par ailleurs, étant donné le volume de transactions et la liquidité offerte par les marchés américains, il est peu probable que la fourchette de cotation aux États-Unis soit plus importante que celle en France.

Existerait-il une concentration des prix plus forte sur les places boursières américaines que sur le marché français ? Étant donné que les actions américaines étaient cotées sur une échelle ("*tick size*") en $\frac{1}{8}$ dollar jusqu'à janvier

2001, il est possible qu'en février 2001 (notre période d'observation), les investisseurs aient gardé une certaine habitude en émettant les ordres aux prix entiers. Une légère auto-corrélation de s pourrait donc être engendrée par la discontinuité des cours successifs.

2. Une explication plus plausible se trouve dans la volatilité temporelle de s . Si celle-ci reste instable même en adoptant une fenêtre d'observation de 512 rentabilités, autrement dit, s'il existe les "**micro bouffées de volatilité**" qui sont tellement fréquentes que même les rentabilités à proximité sont de taille autocorrélée, alors nous devons retrouver ces "**micro bouffées de volatilité**" dans s même après le processus d'uniformisation progressive.

Pour vérifier cette hypothèse, chaque série de rentabilités dans Base I a été découpée en sous-suites de 512 observations, et pour chaque sous-suite, il a été calculée une auto-corrélation des valeurs absolues à l'ordre 1 ; on constate une moyenne de 0,302 pour le coefficient d'auto-corrélation, avec un écart-type de 0,092 pour les actions cotées aux États-Unis, et une moyenne de 0,188 avec un écart-type de 0,085 pour celles en France.

Il semblerait que les données américaines présentent des "**micro bouffées de volatilité**" plus évidentes que pour les données françaises. Ce résultat soutient donc l'explication de la différence de compressibilité par l'auto-corrélation à très court terme des volatilités "*tick-by-tick*".

Les analyses ci-dessus montrent clairement que la différence de compressibilité entre la France et les États-Unis n'implique pas nécessairement un degré d'efficacité moins élevé des marchés américains. Certaines régularités pouvant engendrer des compressions n'implique pas toujours la possibilité de battre la "*buy and hold*".

Conclusion

La complexité de Kolmogorov d'une série donnée mesure sa prévisibilité. L'objectif de ce chapitre était de proposer une méthode générale pour estimer la complexité de Kolmogorov des séries de rentabilités boursières.

Cette méthode, nommée méthode algorithmique, est organisée en 3 étapes successives : 1) transformer les séries de rentabilités en suites de nombres entiers 2) effacer les régularités connues dans les séries de rentabilités, notamment les lois de distribution non-uniformes et les bouffées de volatilité 3) compresser les séries d'entiers avec des algorithmes de compression sans perte. Le taux de compression maximal obtenu pour chaque série est considéré comme l'indicateur de sa prévisibilité.

Cette méthode s'applique aux données de toute fréquence. Par exemple, un degré d'aléa très élevé est constaté dans les rentabilités journalières du DJ dont le taux de compression maximal s'élève à 0,47% après les trois étapes évoquées précédemment. Ce résultat rejoint la plupart des études statistiques sur les rendements successifs des actifs financiers : même si les rentabilités historiques ne sont pas dirigées par le pur hasard, leur prévisibilité reste extrêmement limitée, et par conséquent, battre la “*buy and hold*” demeure, de notre point de vue, quasiment impossible.

En revanche, les régularités sont plus manifestes dans les rentabilités “*tick-by-tick*”. Les 51 séries observées sur le NYSE et le NASDAQ vérifient en moyenne un taux de compression de 4,63%, alors que les 60 observations d'Euronext sont en moyenne compressibles à 0,63%. Ces taux de compression nettement plus élevés que celui des rentabilités journalières, indiquent non seulement la compressibilité des rentabilités intra-journalières, mais aussi les différences de compressibilité entre les 3 places boursières analysées.

À l'aide d'outils statistiques, la stabilité des degrés d'aléa a été vérifiée au sein de chaque pays de cotation, ainsi que la robustesse de la différence de compressibilité entre la France et les États-Unis.

Si les résultats d'observations sont robustes, l'explication théorique du phénomène reste ouverte à toute discussion.

Deux hypothèses sont proposées pour expliquer la différence de compressibilité entre la France et les États-Unis.

1. Les ordres émis aux États-Unis sont plus concentrés autour des nombres entiers que ceux émis en France.
2. Il existe des “**micro bouffées de volatilité**” qui sont plus prononcées aux États-Unis qu'en France.

Ici, nous privilégions la deuxième hypothèse car elle est soutenue par une comparaison des auto-corrélations des rendements absolus à très court terme. Or, le rejet final de la première hypothèse ne serait possible qu'après une investigation sur la concentration des ordres pendant nos périodes d'observation.

Pour conclure, la méthode algorithmique proposée dans ce chapitre est un nouvel outil permettant de quantifier et comparer le degré d'aléa des séries financières. Bien que les algorithmes de compression n'aient pas détecté de régularités “révolutionnaires” permettant de battre le marché, l'introduction de l'approche algorithmique élargit sans doute la gamme des structures détectables, et, ce faisant, rend plus robuste notre conclusion sur l'impossibilité du “*market timing*”. De plus, en proposant un indicateur d'aléa pour chaque série étudiée, la méthode algorithmique ouvre de nouvelles perspectives dans les comparaisons des différents

marchés financiers.

Remarquons toutefois que l'explication des résultats obtenus de cette approche nécessite une étude très large de la littérature, car la prévisibilité potentielle d'une série donnée serait la conséquence d'un grand nombre de facteurs explicatifs. Par conséquent, les taux de compression observés n'ont pas toujours une implication directe au regard de l'efficience informationnelle. Pour apporter plus de lumière à ce sujet, il faut aller plus loin dans notre processus itératif d'effacement de régularités, car après avoir supprimé toutes les régularités neutres à la prévisibilité des baisses boursières (comme les bouffées de volatilité), le taux de compression final mesurerait sans ambiguïté la difficulté à battre la *"buy and hold"*. C'est là une tâche difficile, renvoyée à des travaux futurs.

4. UNE DÉFINITION COMPUTATIONNELLE DES SÉRIES DE COURS IMBATTABLES

Résumé

Dans ce dernier chapitre, nous définissons la notion “*battre le marché*” avec un modèle mathématique. Battre la “*buy and hold*” étant la seule implication directement testable de l’efficience informationnelle, une série de rentabilités est dite “imbattable” si aucune stratégie effective ne peut réussir dans ses tentatives de “*market timing*”.

Trois conclusions découlent de ce dernier chapitre.

- En imposant un seuil à la performance des stratégies simulées, on propose un argument en faveur des “*stop loss rules*” : si cette pratique n’améliore pas toujours le rendement des stratégies de type “*market timing*”, elle n’est pas non plus responsable de leur incapacité à battre la “*buy and hold*”.
- En intégrant les stratégies généralisée, i.e., celles qui peuvent modifier la mise à chaque transaction, on démontre que certaines séries imbattables par les stratégies “tout acheter et tout vendre” sont battables par les stratégies généralisées.
- La vitesse à laquelle une stratégie engendre des différences de gains n’a pas d’incidence sur sa performance finale. La vitesse et la performance sont deux critères relativement indépendants pour évaluer une stratégie donnée.

Introduction

Étant donnée une série de cours successifs, peut-on espérer faire mieux que la stratégie “*buy and hold*” en mettant en place une stratégie active ? Cette question intrigue les chercheurs en finance depuis longtemps sans qu’une réponse définitive soit à ce jour connue. D’un côté, les théoriciens financiers traitent cette problématique en formulant des raisonnements économiques ou psychologiques (Samuelson 1965, Fama 1970, Daniel, Hirshleifer, et Subrahmanyam 1998). D’un autre côté, les économètres (Cowles 1933, Granger 1986, Cooper, Gutierrez, et Hameed 2004), mathématiciens (Bachelier 1900, Mandelbrot 1965) et les économphysiciens (Potter et Bouchaud 1997, Matteo 2007) sont à la recherche de preuves empiriques pour confirmer ou falsifier les théories proposées.

Dans le premier chapitre, nous avons défendu l’idée que la supériorité de la stratégie “*buy and hold*” est une conséquence de la loi de l’offre et de la demande, compatible avec les principaux courants théoriques en finance. Comme toutes les propositions scientifiques, ce point de vue peut aussi être soumis aux tests empiriques. Par exemple, notre proposition serait réfutée, si, en s’appuyant sur des facteurs psychologiques, on trouvait une stratégie de “*market timing*” qui battrait la “*buy and hold*”. Une stratégie sur un cours consiste à définir quand et combien acheter (ou vendre) de titres à ce prix. On fait l’hypothèse que la stratégie en question ne “triche” pas, c’est-à-dire qu’elle ne s’appuie pour décider en “*t*” que sur les informations disponibles à cette date.

La recherche de formules hautement profitables en bourse a toujours été un domaine de travail très actif. Or, malgré de nombreuses études dédiées à ce sujet, aucun résultat concluant n’a été publié. Aussitôt qu’une étude montre l’efficacité d’une stratégie donnée, des critiques méthodologiques sont mises en avant pour défendre la performance du marché.

L’une des explications à cette incohérence méthodologique tient au fait que le terme “battre le marché” n’a jamais fait l’objet d’une définition universelle et précise à notre connaissance. Un test empirique qui, pour certains auteurs, rejette incontestablement la théorie de l’efficacité, peut manquer de robustesse pour d’autres.

Dans ce chapitre, nous essayons de définir la notion de “cours imbattables” dans un langage mathématique, en espérant que l’universalité et la précision de la terminologie mathématique pourront :

1. conduire à un critère commun pour la comparaison des performances des différentes stratégies d’investissement.
2. généraliser la notion des suites aléatoires composées de 0 et de 1 aux suites

de nombres entiers.

3. faciliter la compréhension interdisciplinaire des théories financières.

Ce chapitre apporte donc une définition du concept de “suite de cours imbattable” utile en finance pour donner un sens rigoureux à des expressions communes comme “battre le marché”, ou battre la stratégie “*buy and hold*”.

Par ailleurs, nous démontrons l’existence de telles suites dans différents cadres dont :

- Celui de suites de 0 et de 1 qui sont un cas particulier (certes limité d’un point de vue pratique) de séries de rentabilités.
- Ce cadre est étendu à des séries de rentabilités distribuées normalement. (voir section 4.1.6)
- Nous n’extrapolons pas ce résultat à des séries à la loi de distribution non stable ce qui aurait constitué un cadre probablement plus proche de celui des rentabilités réelles. Toutefois, cette limite n’invalide pas la définition proposée et ce cadre particulier pourrait être exploré dans des travaux ultérieurs.

Ce travail entretient cependant un rapport au réel certain au travers de l’exploration de stratégies génériques en finance : celles des “*stop loss rules*” (Kaminski et Lo 2007) et celles dites géométriques (Ross 1999).

En synthèse, la contribution se veut ici principalement théorique, interdisciplinaire ouvrant de nouvelles pistes de recherche à la frontière entre théorie financière et théorie algorithmique de la complexité.

Inspirée par la définition de suites aléatoires au sens de Schnorr (1971) (présentées dans la section 3.3, page 104), on définit les séries de cours imbattables par celles qui ne sont battues par aucune stratégie effective. Ici, une série de cours est dite “battue” par une stratégie, $d(\cdot)$, si $d(\cdot)$ peut engendrer infiniment plus de profits que la “*buy and hold*”.

Au lieu de comparer des stratégies selon leur espérance de gains, comme ce qui est souvent réalisé en finance, notre définition de cours imbattables s’inscrit dans un nouveau cadre théorique : celui de la théorie de la calculabilité.

En effet, traditionnellement, une série financière est toujours présentée dans le cadre des probabilités. Par exemple, en supposant que la série de rentabilités x_0, x_1, \dots, x_n est réalisée par une collection de variables aléatoires $i.i.d.N(0, 1)$, on déduit les propriétés de x_0, x_1, \dots, x_n en étudiant cette dernière collection de variables. Comme celle-ci peut engendrer un nombre infini de suites différentes, en réalité, on analyse toujours **une famille** de séries probables mais non une suite spécifique. L’étude d’un échantillon réel dans le cadre des probabilités

consiste donc à identifier “la population” à laquelle celui-ci peut bien appartenir. Les conclusions obtenues sont toujours de type “ x_0, x_1, \dots, x_n est réalisée par une collection de variables aléatoires $i.i.d.N(0, 1)$ qui est imbattable”.

En revanche, dans le cadre de la calculabilité, on peut étudier directement une série spécifique, et formuler des conclusions comme “la série x_0, x_1, \dots, x_n est imbattable”. C’est justement cette nouveauté conceptuelle qui nous permet d’appliquer les outils algorithmiques sur des séries financières.

Dans ce chapitre, nous proposons trois notions de “cours imbattables” en étudiant des modèles de stratégies de plus en plus proches des conditions d’échanges prévalant sur les marchés financiers réels :

- Un premier modèle de stratégie “**tout acheter et tout vendre**” est proposé dans la section 4.1.

Deux résultats intéressants sont obtenus en utilisant ce modèle simplifié :

1. Battre la “*buy and hold*” revient à identifier les baisses du marché (Brandouy, Mathieu, et Veryzhenko 2009). La différence de gains entre une stratégie et la “*buy and hold*” se trouve donc dans les moments où cette stratégie impose de sortir du marché.
 2. En introduisant une politique de blocage à un euro, nous apportons un support théorique aux “*stop loss rules*” pratiquées par certains fonds d’investissement : une stratégie qui bat le marché à long terme, ne sera pas influencée par les ruptures temporaires imposées par cette politique de blocage.
- Bien que les stratégies “tout acheter tout vendre” offrent un cadre d’analyse convenable, en réalité, les investisseurs varient leurs “mises” selon leur niveau de confiance et l’état de leur portefeuille. C’est pourquoi on propose, dans la section 4.4, un modèle de stratégie généralisée qui peut effectuer un choix libre concernant la proportion de richesse engagée dans chaque intervention sur le marché. Puis, en comparant les deux modèles proposés, nous montrons formellement que les stratégies généralisées peuvent battre plus de séries financières que celles du type “tout acheter tout vendre”.
 - Une bonne stratégie doit non seulement engendrer des profits satisfaisants, mais aussi être rapide dans la réalisation de gains. Pour prendre en compte ce deuxième critère, nous définissons dans la section 4.6, la vitesse à laquelle une stratégie bat la “*buy and hold*”.

4.1 Modèle “tout acheter et tout vendre”

Pour définir les cours imbattables, il faut commencer par élucider la notion de “battre le marché”. Étant donnée une série de cours boursiers, quelle rentabilité une stratégie doit engendrer pour être considérée comme ayant battu le marché ? Une première idée consiste à exprimer directement une différence de gains.

- Soit $c_0, c_1, c_2, \dots, c_n, \dots$ ($c_n \in \mathbb{N}^*, n \in \mathbb{N}$), la série de cours boursiers avec c_n représentant le “nième” cours ¹.
- Soit $l_n \in \mathbb{Q}^+$ une fonction qui indique la liquidité disponible à l’étape “ n ”. Ici, on impose $l_0 \in \mathbb{N}^*$, ce qui signifie un capital initial positif pour toutes les stratégies considérées.
- Soit a_n la fonction qui calcule le nombre de titres disponible à l’instant “ n ”. Sans perdre de généralité, on impose $a_0 = 0$.
- Soit $d(c_0, c_1, c_2, \dots, c_{n-1})$ une stratégie d’investissement satisfaisant :

$$d(c_0, c_1, c_2, \dots, c_{n-1}) \in \{0, -1, +1\}$$

Ici,

- “ -1 ” signifie vendre tous les titres du portefeuille,
- “ $+1$ ” acheter avec toute la liquidité disponible du moment,
- et “ 0 ” ne rien faire.
- $d(c_0, c_1, c_2, \dots, c_{n-1}) = -1$ seulement si $a_{n-1} > 0$. Autrement dit, il est impossible de vendre une action avant de l’acheter (“*short selling*”).

Pour que la stratégie $d(\cdot)$ puisse faire l’objet d’une simulation computationnelle, nous imposons que celle-ci soit calculable (au sens de la théorie de la calculabilité présentée dans l’annexe K).

$d(\cdot)$, l_n et a_n ainsi définis vérifient les relations suivantes :

- Toutes les stratégies commencent par un capital initial en liquide, autrement dit, $l_0 \in \mathbb{N}^*, a_0 = 0$.
- À l’instant 1, $d(c_0)$ prend sa décision selon c_0 . Comme le capital initial est toujours en liquide, la première décision de $d(\cdot)$ ne peut qu’être “acheter” ou “ne rien faire”. C’est-à-dire que,

$$l_1 = l_0 - d(c_0)l_0, \quad a_1 = a_0 + d(c_0)\frac{l_0}{c_0} \quad d(c_0) \in \{0, 1\} \quad (4.1)$$

- À l’instant 2, $d(c_0, c_1)$ (désormais d_2) peut être $-1, 0$ ou 1 , on a,

$$\begin{cases} l_2 = a_1 c_1, & a_2 = 0 & \text{si } d_2 = \{-1\} \\ l_2 = l_1 - d_2 l_1, & a_2 = a_1 + d_2 \frac{l_1}{c_1} & \text{si } d_2 = \{0, +1\} \end{cases} \quad (4.2)$$

1. Remarquons que selon cette définition, toute suite de nombres entiers positifs peut être considérée comme une série de cours, sa série de rentabilités correspondante est donc définie sur l’ensemble des nombres réels. Plus particulièrement, une suite composée de 0 et de 1 peut aussi être considérée comme une série de rentabilités, bien que cela ne soit pas le cas sur les marchés financiers réels.

– À l’instant n , $d(c_0, c_1, \dots, c_{n-1})$ (désormais d_n), l_n et a_n vérifient :

$$\begin{cases} l_n = a_{n-1}c_{n-1}, & a_n = 0 & \text{si } d_n = \{-1\} \\ l_n = l_{n-1} - d_n l_{n-1}, & a_n = a_{n-1} + d_n \frac{l_{n-1}}{c_{n-1}} & \text{si } d_n = \{0, +1\} \end{cases} \quad (4.3)$$

À l’instant n , la richesse engendrée par $d(\cdot)$ (notée r_n) s’élève à :

$$r_n = l_n + a_n \times c_n \quad (4.4)$$

Ici, on remarque que

- $r_n \in \mathbb{Q}$ pour chaque instant n . En effet, par définition, $c_0, c_1, \dots, c_n, \dots$ est une suite de nombres entiers, la richesse initiale $l_0 \in \mathbb{N}^*$, et $\frac{l_{n-1}}{c_{n-1}} \in \mathbb{Q}$. Comme la fonction de décision $d(\cdot) \in \{-1, 0, 1\}$, et $l_n, a_n \in \mathbb{Q}$, la fonction de richesse $r_n \in \mathbb{Q}$ également.
- Étant donnée une série de cours $c_0, c_1, \dots, c_{n-1}, \dots$, la fonction r_n est calculable, si $d(\cdot)$ l’est également.

En même temps, la stratégie de référence (“*buy-and-hold*”) qui commence par le même capital initial l_0 , engendrera à l’instant “ n ” une richesse de

$$r_n^* = \frac{c_n}{c_0} l_0$$

En comparant la performance de $d(\cdot)$ et celle de la “*buy and hold*”, on peut définir

la notion de “battre le marché” comme suit :

Définition “Battre le marché” 4.1.0.1

1. Une série de cours $c_0, c_1, c_2, \dots, c_n, \dots$ est “faiblement battue” par la stratégie d’investissement $d(\cdot)$ **sii** :

$$\limsup_{n \rightarrow \infty} (r_n / r_n^*) = +\infty \quad (4.5)$$

2. $c_0, c_1, c_2, \dots, c_n, \dots$ est “fortement battu” par la stratégie d’investissement $d(\cdot)$ **sii**² :

$$\liminf_{n \rightarrow \infty} (r_n / r_n^*) = +\infty \quad (4.6)$$

Afin de bien distinguer les deux définitions proposées ci-dessus, nous estimons qu’il est utile de faire un point sur la définition de la limite supérieure (désormais “lim sup”) et la limite inférieure (désormais “lim inf”).

Définition “Limite inf (resp. sup)” 4.1.0.2

$$\liminf_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \left(\min_{m \geq n} x_m \right).$$

$$\text{resp. } \limsup_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \left(\max_{m \geq n} x_m \right).$$

Les propriétés suivantes découlent directement de la définition 4.1.0.2. Comme elles seront souvent utilisées dans la suite de ce chapitre, nous estimons qu’il est nécessaire de les rappeler formellement.

Rappels 4.1.0.3

$$\limsup x_n = a \in \mathbb{R} \Leftrightarrow$$

$$\forall \varepsilon > 0, \forall m \in \mathbb{N}, \exists n > m, x_n > a - \varepsilon$$

$$\text{et } \forall \varepsilon > 0, \exists m \in \mathbb{N}, \forall n > m, x_n < a + \varepsilon$$

$$\liminf x_n = a \in \mathbb{R} \Leftrightarrow$$

$$\forall \varepsilon > 0, \forall m \in \mathbb{N}, \exists n > m, x_n < a + \varepsilon$$

$$\text{et } \forall \varepsilon > 0, \exists m \in \mathbb{N}, \forall n > m, x_n > a - \varepsilon$$

En ce qui nous concerne, $\limsup_{n \rightarrow \infty} r_n / r_n^* = \infty$ signifie que, pour toutes valeurs $M \in \mathbb{R}$ et $m \in \mathbb{N}$, il existe toujours un instant $n > m$ où la différence de gains r_n / r_n^* dépasse M .

4.1. Modèle “tout acheter et tout vendre”

En revanche, $\liminf_{n \rightarrow \infty} r_n/r_n^* = \infty$ signifie que pour tout $M > 0$, il existe un instant $m \in \mathbb{N}$ qui satisfait $r_n/r_n^* > M$ pour tout $n > m$.

Autrement dit, quand une stratégie “bat faiblement” une série de cours donnée, il faut toujours choisir le “bon moment” pour sortir du marché avec une différence de gains satisfaisante. Alors que dans le cas “bat fortement”, il suffit de rester suffisamment longtemps sur le marché pour réaliser une différence de gains importante.

Étant donnée l'équation 4.7, une stratégie battant fortement une série de cours donnée la bat faiblement aussi. Cette implication a un effet sur la notion de cours imbattables.

$$\liminf_{n \rightarrow \infty} r_n/r_n^* = \infty \Rightarrow \limsup_{n \rightarrow \infty} r_n/r_n^* = \infty \quad (4.7)$$

A partir de la notion de “battre le marché”, nous pouvons formuler une définition de cours imbattables.

Définition “Cours imbattables” 4.1.0.4

Une série de cours est dite “**faiblement** (resp. **fortement**) **imbattable**”, si aucune stratégie d'investissement ne peut la battre **fortement** (resp. **faiblement**).

Une suite de rentabilités est “faiblement (resp. fortement) imbattable”, si son cours correspondant³ est “faiblement (resp. fortement) imbattable”.

De façon contre-intuitive, selon la définition 4.1.0.4, les cours **faiblement** imbattables ne correspondent pas à ceux qui ne sont **faiblement** battus par aucune stratégie d'investissement. Ici, nous avons “inversé” les termes en associant “**faiblement** imbattable” aux suites qui ne sont **fortement** battues par aucune stratégie pour faire en sorte que la notion “faiblement imbattable” soit moins exigeante que celle de “fortement imbattable”.

Comme un cours fortement battable est toujours faiblement battable, grâce à l'inversion des termes, la notion “fortement imbattable” implique également “faiblement imbattable”.

Venons-en aux propriétés des cours imbattables.

4.1.1 Différence de gains en termes de rentabilités

La différence de gains (r_n/r_n^* dans la définition 4.1.0.1) peut aussi s’exprimer en termes de rentabilités.

Soit $c_0, c_1, \dots, c_n, \dots$, une série de cours entiers. $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, la suite des rentabilités lui correspondant, alors, on peut définir :

Définition “Sous-suite raisonnablement extraite” 4.1.1.1

Une sous-suite $\sigma_0^s, \sigma_1^s, \dots, \sigma_n^s, \dots$ est dite “raisonnablement extraite” de la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, s’il existe une fonction calculable, notée $s(c_0, c_1, \dots, c_{n-1})$, qui, à chaque instant n , décide si la prochaine rentabilité, σ_n se trouve dans la sous-suite extraite.

Appliquer une stratégie $d(\cdot)$ sur une série de cours c_0, c_1, \dots, c_n , revient à diviser la suite des rentabilités lui correspondant, $\sigma_0, \sigma_1, \dots, \sigma_n$ en deux sous-suites raisonnablement extraites : celle des rentabilités **utilisées**⁴ par $d(\cdot)$ et celle des rentabilités **refusées** par $d(\cdot)$, notée $\sigma_0^s, \sigma_1^s, \dots, \sigma_n^s$.

Proposition 4.1.1.2

La différence de gains entre $d(\cdot)$ et la “buy and hold” est donnée par l’équation (4.8) :

$$r_n/r_n^* = 2^{\sum_{i=1}^n -\sigma_i^s} \tag{4.8}$$

Démonstration : Comme le nombre d’actions détenu à l’instant n (c’est-à-dire la fonction a_n définie dans la section 4.1, page 182) dépend de la stratégie $d(c_0, c_1, \dots, c_{n-1})$, a_n est calculable **si** $d(\cdot)$ l’est également.

Définissons la fonction d’extraction $s(c_0, c_1, \dots, c_{n-1})$ (désormais s_n) comme suit :

$$s_n = \begin{cases} 0 & \text{si } a_n = 0 \\ +1 & \text{sinon} \end{cases} \tag{4.9}$$

Ici, $s_n = 0$ signifie que $\sigma_n \in \sigma_0^s, \sigma_1^s, \dots, \sigma_n^s$, et $s_n = 1$ le cas contraire. De toute évidence, s_n est calculable si $d(\cdot)$ l’est également.

La richesse engendrée par $d(\cdot)$ à l’instant n peut ainsi être calculée par l’équation (4.10) :

$$r_n = l_0 2^{s_1 \sigma_1 + s_2 \sigma_2 + \dots + s_n \sigma_n} \tag{4.10}$$

Alors que la richesse de la “buy and hold” est :

$$r_n^* = l_0 2^{\sigma_1 + \sigma_2 + \dots + \sigma_n} \tag{4.11}$$

4. Au sens où $d(\cdot)$ décide de tenir des titres à l’instant n . La rentabilité σ_n contribue aux gains accumulés de $d(\cdot)$ si elle est utilisée.

La différence de gains,

$$r_n/r_n^* = 2^{s_1\sigma_1+s_2\sigma_2+\dots+s_n\sigma_n-(\sigma_1+\sigma_2+\dots+\sigma_n)} = 2^{\sum_{i=1}^n -\sigma_i^s} \quad (4.12) \quad \blacksquare$$

À partir de cette relation étroite entre $d(\cdot)$ et $s(\cdot)$, nous pouvons définir les suites des rentabilités imbattables par extraction.

Définition “Battre par extraction” 4.1.1.3

Une suite de rentabilités $\sigma_1, \sigma_2, \dots, \sigma_n, \dots$, est dite “faiblement (resp. fortement) battable” par les méthodes d’extraction, s’il existe une sous-suite raisonnablement extraite (désormais sous-suite raisonnable), notée $\sigma_1^s, \sigma_2^s, \dots, \sigma_n^s, \dots$, qui vérifie :

$$\liminf_{n \rightarrow \infty} \sum_{i=1}^n \sigma_i^s = -\infty \quad (4.13)$$

resp. :

$$\limsup_{n \rightarrow \infty} \sum_{i=1}^n \sigma_i^s = -\infty \quad (4.14)$$

On obtient ainsi la définition de suites imbattables par extraction :

Définition “Imbattable par extraction” 4.1.1.4

Une série de cours est dite “fortement (resp. faiblement) imbattable”, si aucune fonction d’extraction raisonnable ne peut la battre faiblement (resp. fortement).

Les suites fortement battables par des méthodes d’extraction existent. Par exemple, la suite des rentabilités $10, -1, 1, -1, 1, -1, 1 \dots$ est fortement battable, car la fonction calculable qui retient uniquement les termes pairs, extraira une sous-suite raisonnable dont la somme tend vers $-\infty$.

Les suites fortement imbattables existent également. Par exemple, les suites de rentabilités positives sont fortement imbattables : pour extraire une sous-suite infinie dont la somme tend vers $-\infty$, il faut que les termes négatifs existent dans la suite d’origine. Les suites positives qui ne présentent aucune baisse ne peuvent donc dans aucun cas être battables par des stratégies d’investissement. Cette idée sera développée dans la section suivante.

4.1.2 Propriété partagée par les cours battables

Nous proposons maintenant, une condition nécessaire pour qu’une suite soit faiblement batable par des stratégies d’investissement :

Soit $c_0, c_1, \dots, c_n, \dots$ une suite de cours faiblement batable, et $\sigma_1, \sigma_2, \dots, \sigma_n, \dots$ la suite des rentabilités lui correspondant.

Proposition 4.1.2.1

Alors, la suite $\sigma_1, \sigma_2, \dots, \sigma_n, \dots$ vérifie :

$$\sum_{\sigma_n < 0}^{\infty} \sigma_n = -\infty \tag{4.15}$$

Démonstration : D’après la proposition 4.1.1.2, la différence de gains entre une stratégie quelconque et la “*buy and hold*” est donnée par :

$$2^{\sum_{i=0}^n -\sigma_i^s}$$

Où les σ_i^s forment une sous-suite raisonnable de $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$.
Ainsi, pour que cette différence de gains puisse avoir une *lim sup* infinie, il faut qu’il existe suffisamment de rentabilités négatives dans $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$.
C’est-à-dire que

$$\sum_{\sigma_n < 0}^{\infty} \sigma_i = -\infty \quad \blacksquare$$

La condition 4.15 implique l’impossibilité de battre des suites positives. Par exemple, la série de cours $1, 2, 4, \dots, 2^n, \dots$ correspondant à la suite des rentabilités $1, 1, 1, \dots, 1, \dots$ est imbattable, car la meilleure façon d’investir sur cette suite consiste à acheter dès le premier instant, et à garder les actifs pour profiter de chaque hausse. Toute “sortie” du marché engendra “un manque à gagner” par rapport à la “*buy and hold*”.

En revanche, la série de cours $1, 2, 1, 2, \dots, 1, 2, \dots$ correspondant à la suite des rentabilités $1, 1, -1, 1, -1, \dots, 1, -1, \dots$ est fortement batable, car la stratégie “*buy and hold*” engendrera un gain, $r_n^* = l_0$ ou $2l_0$ selon la parité de n . Tandis que la stratégie “acheter aux termes pairs et vendre aux termes impairs” engendra une richesse r_n qui tend vers l’infini exponentiellement.

4.1.3 Condition d'existence des cours imbattables

Selon la définition 4.1.0.4 (page 185), la stratégie $d(\cdot)$ bat faiblement (resp. fortement) une série de cours donnée, si et seulement si la $\lim \sup$ (resp. $\lim \inf$) de la différence de gains est $+\infty$.

Pourquoi exigeons-nous ici une différence de gains infinie ? Est-ce une condition trop stricte ? Serait-il possible de la remplacer par, par exemple,

$$\limsup_{n \rightarrow \infty} r_n / r_n^* > 100000$$

Nous montrons dans cette section que c'est la condition $\limsup_{n \rightarrow \infty} r_n / r_n^* = \infty$ qui exclut la possibilité de “battre le marché par chance”, un phénomène souvent associé au “data snooping” dans la littérature en finance ⁵.

Proposition 4.1.3.1

Si la condition requise pour qu'une suite soit “faiblement (resp. fortement) battue” était :

$$\limsup(\text{resp.} \liminf)_{n \rightarrow \infty} (r_n / r_n^*) = D$$

où $D \in \mathbb{R}^+ < \infty$, alors toutes les séries de cours vérifiant la condition 4.15 (page 188) seraient faiblement (resp. fortement) battables.

Démonstration : Soit $\sigma_0, \sigma_1, \dots, \sigma_n$, une suite de rentabilités vérifiant la condition 4.15, c'est-à-dire que $\sum_{i=0, \sigma_i < 0}^{\infty} \sigma_i = -\infty$.

Ainsi, quelle que soit la valeur de D , $\exists m \in \mathbb{N}^*$ vérifiant $\sum_{i=0, \sigma_i < 0}^m \sigma_i < -\log(D)$. Soit c_0, c_1, \dots, c_m le début du cours correspondant aux premières rentabilités $\sigma_0, \dots, \sigma_m$. À partir de ces premiers m cours, nous pouvons établir au maximum 2^m stratégies différentes⁶. Parmi ces 2^m stratégies, on peut certainement en identifier une qui, **par pure chance**, détient des titres quand le cours augmente et du “cash” lorsque celui-ci baisse.

La différence de gains engendrée par cette stratégie chanceuse est donnée par l'équation (4.16) :

$$r_m / r_m^* = 2^{\sum_{i=0, \sigma_i < 0}^m -\sigma_i} \quad (4.16)$$

Par construction, $\sum_{i=0, \sigma_i < 0}^m \sigma_i < -\log(D)$, ce qui implique

$$r_m / r_m^* = 2^{\sum_{i=0, \sigma_i < 0}^m -\sigma_i} > 2^{\log(D)} = D \quad (4.17)$$

5. Le phénomène “data snooping” est l'un des principaux “pièges” à éviter dans la recherche des stratégies gagnantes en finance. Parfois, on constate par pure chance des “régularités non-récurrentes” dans des séries financières. Une stratégie exploitant ce type de régularités peut surperformer le marché pendant une période relativement courte. Or, ces gains temporaires seront annihilés par les futures pertes, car les régularités exploitées disparaîtront dans le temps. Les régularités non-récurrentes ne doivent donc pas changer notre conclusion sur la possibilité de battre une suite donnée.

6. 2^m stratégies possibles, car à chaque instant t , $d(\cdot)$ peut prendre deux parmi les trois valeurs possibles ($-1, 0$ ou 1), selon sa décision prise à l'instant $t - 1$.

Ainsi, quelle que soit la valeur de D , il existe toujours un début de stratégie dont la différence de gains $r_m/r_m^* > D$ pour toute série de cours vérifiant la condition 4.15.

La fonction $d(\cdot)$ qui consiste à suivre la “stratégie chanceuse” avant l’instant m et qui, à partir de l’instant $m + 1$, pratique la “*buy and hold*”, gardera son avance par rapport à la “*buy and hold*” jusqu’à ∞ . ■

Toutefois, remarquons que même en remplaçant la condition $\limsup_{n \rightarrow \infty} (r_n/r_n^*) = +\infty$ par $\limsup_{n \rightarrow \infty} (r_n/r_n^*) = D$ les suites imbattables existeraient toujours. Les suites de rentabilités positives qui ne vérifient pas la condition 4.15, resteront fortement imbattables.

Bien sûr, ceci ne rend pas légitime l’usage de la condition $\limsup_{n \rightarrow \infty} (r_n/r_n^*) = D$, car si nous formulons une définition de suites imbattables, c’est essentiellement pour distinguer, parmi les suites satisfaisant la condition 4.15, celles qui sont imbattables de celles qui ne le sont pas.

4.1.4 L’aspect aléatoire d’une suite batable

Dans cette section, nous allons montrer qu’un cours batable selon la définition 4.1.0.4 peut également avoir des termes aléatoires (i.e. peut contenir une sous-suite aléatoire). Pour battre une série de cours donnée, il suffit d’identifier une seule régularité récurrente, même si celle-ci est peu présente dans la suite $c_0, c_1, \dots, c_n, \dots$.

Pour montrer une telle proposition, nous allons utiliser la notion des suites aléatoires au sens de Martin-Löf (désormais “suites de Martin-Löf”), concept que nous avons rappelé dans le chapitre précédent lors de la présentation du hasard en informatique (cf. page 104).

Une suite de Martin-Löf vérifie, avec une probabilité de 100%, toutes les propriétés que nous pouvons constater dans une suite réalisée par une collection de variables aléatoires *i.i.d.* $U(0, 1)$ ⁷. Par exemple, les propriétés suivantes sont souvent utilisées dans ce chapitre :

1. Toute sous-suite raisonnablement extraite d’une suite de Martin-Löf reste une suite de Martin-Löf.
2. Toute suite de Martin-Löf a une fréquence limite de 0,5 pour le terme 1.

7. Dans ce chapitre, sauf précision particulière, l’expression $U(0, 1)$ désigne la loi suivie par la variable $X \in \{0, 1\}$, $Prob(X = 1) = 0,5$ et $Prob(X = 0) = 0,5$.

3. Soit $\sigma_0, \sigma_1, \dots, \sigma_n$ une suite de Martin-Löf, alors la valeur absolue, $|\sum_{i=0}^n \sigma_i - \frac{0 \times n + 1 \times n}{2}|$ croît à la vitesse \sqrt{n} , ce qui implique que

$$\limsup_{n \rightarrow \infty} (\text{resp. inf}) \sum_{i=0}^n \sigma_i - \frac{1 \times n + 0 \times n}{2} = \infty (\text{resp. } -\infty). \quad (4.18)$$

4. Une collection de variables aléatoires *i.i.d.* $U(0, 1)$ engendre des suites de Martin-Löf avec une probabilité de 100%.
5. ...

Après avoir rappelé brièvement les propriétés des suites de Martin-Löf, nous allons maintenant construire une suite faiblement battable qui contient les termes aléatoires dans le sens de Martin-Löf.

Soit :

- $a_0, a_1, \dots, a_n, \dots$ une suite de Martin-Löf dont les 0 sont remplacés par -1 . Sans perdre de généralité, nous imposons $a_0 = 1$.
- $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ la suite de rentabilités obtenue en remplaçant les a_{10m+1} ($m \in \mathbb{N}$) par 10 comme illustré par les deux premières lignes du tableau 4.1.
- $c_0, c_1, \dots, c_n, \dots$ la série de cours correspondant à $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$.

Instant	0	1	2	3	4	5	6	7	8	9	10	11	12	...
a_0, \dots, a_n, \dots	1	-1	1	1	-1	1	1	-1	-1	-1	1	-1	1	...
↓	1	10										10		...
$\sigma_0, \dots, \sigma_n, \dots$	1	10	1	1	-1	1	1	-1	-1	-1	1	10	1	...
$r_n^* = c_n$	1	2^{10}	2^{11}	2^{12}	2^{11}	2^{12}	2^{13}	2^{12}	2^{11}	2^{10}	2^{11}	2^{21}	2^{22}	...
r_n	1	2^{10}	2^{10}	2^{10}	2^{10}	2^{10}	2^{10}	2^{10}	2^{10}	2^{10}	2^{10}	2^{20}	2^{20}	...
r_n/r_n^*	1	1	2^{-1}	2^{-2}	2^{-1}	2^{-2}	2^{-3}	2^{-2}	2^{-1}	1	2^{-1}	2^{-1}	2^{-2}	...

Tab. 4.1: Construction de la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ et le calcul de la différence de gains r_n/r_n^* .

Proposition 4.1.4.1

La suite des rentabilités ainsi construite (notée $\sigma_0, \sigma_1, \dots, \sigma_n$) correspond à une série de cours entiers.

Démonstration : Par définition, $c_{10m+1} = c_{10m} 2^{\sigma_{10m+1}}$. Comme $\sigma_{10m+1} = 10$, on obtient $c_{10m+1} = c_{10m} 2^{10}$. Étant donné $\sigma_{10m+2} + \sigma_{10m+3} + \dots + \sigma_{10m+10} \geq -9$, $c_{10(m+1)+1} \geq c_{10m+1} 2^{-9} = c_{10m} 2^{10-9} = 2c_{10m}$. C’est-à-dire que

$$c_{10m+1}, c_{10m+2}, \dots, c_{10m+10} \in \mathbb{N}^* \text{ si } c_0 \in \mathbb{N}^*.$$

Comme nous avons imposé $c_0 = a_0 = 1$, la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n$ correspond bien à une série de cours entiers. ■

Fondée sur la suite $\sigma_0, \sigma_1, \dots, \sigma_n$, nous pouvons démontrer la proposition suivante :

Proposition 4.1.4.2

La stratégie $d(c_0, c_1, \dots, c_{n-1})$ qui utilise uniquement les $(10m + 1)$ ièmes termes de la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ bat faiblement la série de cours $c_0, c_1, \dots, c_n, \dots$.

Démonstration : Selon la proposition 4.1.1.2 (page 186), la différence de gains entre $d(\cdot)$ et la “buy and hold” dépend de la somme des rentabilités refusées. Dans notre cas, la différence de gains est donnée par :

$$2^{\sum_{i=1, i \neq 10m+1}^n -\sigma_i} \quad m \in \mathbb{N}$$

Par construction, les rentabilités refusées, notées $\sigma_0^r, \sigma_1^r, \dots, \sigma_n^r, \dots$, forment une suite de “-1” et “1” aléatoire dans le sens de Martin-Löf. En appliquant la troisième propriété des suites de Martin-Löf rappelée à la page 190, on obtient,

$$\limsup_{n \rightarrow \infty} (\text{resp. inf}) \sum_{i=1}^n -\sigma_i^r - \frac{-1 \times n + 1 \times n}{2} = \infty (\text{resp. } -\infty)$$

C'est-à-dire que,

$$\limsup_{n \rightarrow \infty} r_n / r_n^* = \limsup_{n \rightarrow \infty} \sum_{i=1, i \neq 10m+1}^n -\sigma_i = \infty$$

Donc, $d(\cdot)$ bat faiblement la série de cours $c_0, c_1, \dots, c_n, \dots$. ■

Remarquons que dans notre exemple, la suite $\sigma_1, \sigma_2, \dots, \sigma_n \dots$ est dominée par un élément aléatoire : $\frac{9}{10}$ de ses termes varient de façon aléatoire au sens de Martin-Löf. La prévisibilité de $\frac{1}{10}$ des termes s'est avérée suffisante pour faiblement battre la suite $\sigma_1, \sigma_2, \dots, \sigma_n \dots$.

Maintenant, il est naturel de nous demander si les séries de cours fortement battables peuvent également contenir des termes aléatoires.

Proposition 4.1.4.3

Certaines suites de rentabilités sont **fortement** battables tout en contenant des sous-suites aléatoires.

Démonstration : Commençons toujours par la suite a_0, \dots, a_n, \dots composée de -1 et 1 aléatoire au sens de Martin-Löf (1966). Cette fois-ci, pour construire la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n$, nous remplaçons les a_{10m+1} par 20 et les a_{10m+2} par -10. Comme dans le cas précédent, nous imposons $\sigma_0 = 1$. Ce

4.1. Modèle “tout acheter et tout vendre”

	0	1	2	3	4	5	6	7	8	9	10	11	12	...
a_0, \dots, a_n, \dots	1	-1	1	1	-1	1	1	-1	-1	-1	1	-1	1	...
↓	1	20	-10									20	-10	...
$\sigma_0, \dots, \sigma_n, \dots$	1	20	-10	1	-1	1	1	-1	-1	-1	1	20	-10	...
$r_n^* = c_n$	1	2^{20}	2^{10}	2^{11}	2^{10}	2^{11}	2^{11}	2^{10}	2^9	2^8	2^9	2^{29}	2^{19}	...
r_n	1	2^{20}	2^{20}	2^{20}	2^{20}	2^{20}	2^{20}	2^{20}	2^{20}	2^{20}	2^{20}	2^{40}	2^{40}	...
r_n/r_n^*	1	1	2^{10}	2^9	2^{10}	2^9	2^9	2^{10}	2^{11}	2^{12}	2^{11}	2^{11}	2^{21}	...

Tab. 4.2: Construction de la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n$ et le calcul de la différence de gains r_n/r_n^* . Ici, la différence de gains dépend de la somme des rentabilités refusées $r_n/r_n^* = \sum_{m=1}^n \sum_{i=10m+2}^{10m+10} (-\sigma_i)$. Comme $\sigma_{10m+2} = -10$ et $\sum_{i=3}^{10} \sigma_{10m+i} \geq 8$. La différence de gains est au moins multipliée par 2^2 à chaque fois que m augmente de 1. r_n/r_n^* tend donc vers ∞ quand $b \rightarrow \infty$.

processus de construction est décrit dans le tableau 4.2.

Soit $\sigma_0, \sigma_2, \dots, \sigma_n$, la suite des rentabilités ainsi construite, $c_0, c_1, \dots, c_n, \dots$ la série de cours correspondant à $\sigma_0, \sigma_2, \dots, \sigma_n$, et $d(c_0, c_1, \dots, c_{n-1})$ une stratégie calculable qui utilise uniquement les σ_{10m+1} . Nous allons montrer que $d(\cdot)$ bat fortement $c_0, c_1, \dots, c_n, \dots$.

Toujours selon la proposition 4.1.1.2 (page 186), la différence de gains engendrée par $s(\cdot)$ s'exprime comme suit :

$$r_n/r_n^* = 2^{\sum_{i=1, i \neq 10m+1}^n -\sigma_i} \quad m \in \mathbb{N}$$

Comme

1. $\sigma_{10m+2} = -10$
2. et $\sum_{i=10m+3}^{10m+10} -\sigma_i \geq -8$

nous avons $\sum_{i=10m+2}^{10m+10} -\sigma_i \geq 10 - 8 = 2$ pour tout $m \in \mathbb{N}$.

Ainsi, à chaque fois m augmente de 1, r_n/r_n^* doit au moins augmenter de 2^2 .

Nous avons donc,

$$\liminf_{n \rightarrow \infty} r_n/r_n^* = \liminf_{n \rightarrow \infty} 2^{\sum_{i=1, i \neq 10m+1}^n -\sigma_i} = \infty$$

$d(\cdot)$ est une stratégie calculable qui peut engendrer une différence de gains infinie sur le cours $c_0, c_1, \dots, c_n, \dots$ qui contient des sous-suites aléatoires. ■

Il est convenable de préciser ici que pour n'exploiter qu'une partie des rentabilités comme nous l'avons fait dans l'exemple précédent, mieux vaut respecter une règle d'or : “**ne rien faire dans l'incertitude**”. Sinon, on risquerait de perdre les gains sûrs mais modestes à cause des pertes éventuellement conséquentes.

Bien sûr, dans l'exemple précédent, la stratégie $d(\cdot)$ peut toujours battre la suite des rentabilités $\sigma_0, \sigma_2, \dots, \sigma_n$, même en utilisant une partie de ses termes

aléatoires. Cela est dû à la faible volatilité des termes aléatoires par rapport aux gains certains qu’on peut engendrer sur les σ_{10m+1} .

La règle d’or est malgré tout conseillée dans ce cas de figure, car la volatilité des termes aléatoires peut s’avérer plus importante pour d’autres séries de rentabilités et la conséquence des interventions “risquées” serait plus grave que dans l’exemple donné.

En revanche, la règle d’or simple à comprendre va souvent à l’encontre de la nature humaine, et ce, même pour les traders les plus expérimentés. Comme Lefèvre (1923) l’a formulé dans son autobiographie :

“In fact, I always made money when I was sure I was right before I began. What beat me was not having brains enough to stick to my own game that is, to play the market only when I was satisfied that precedents favored my play.” (Lefèvre 1923, p.185)

4.1.5 Différence théorique entre “faiblement battre” et “fortement battre”

Dans la définition 4.1.0.4 (page 185), nous distinguons deux niveaux de hasard : les suites “faiblement imbattables” et celles qui sont “fortement imbattables”. Quelle est la différence entre ces deux définitions ? Puisque cette distinction prend un sens uniquement s’il existe des suites faiblement mais non fortement battables, nous allons donc comparer ces deux notions en proposant des exemples concrets à fins d’illustrations.

À priori, la différence entre la limite sup et la limite inf d’une suite donnée peut être ∞ . Autrement dit, la différence de gains (entre une stratégie donnée et la “buy and hold”) peut avoir une limite sup infinie sans que sa limite inf le soit aussi. La différence entre “faiblement imbattable” et “fortement imbattable” existe donc au moins *à priori*.

Pour démontrer l’existence d’une telle différence, nous allons construire une suite de rentabilités faiblement mais non fortement battable par des stratégies d’investissement.

Soit :

- $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ une suite composée de 1 et de -1 aléatoire dans le sens de Martin-Löf.
- $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ la suite des rentabilités obtenue en remplaçant tous les termes pairs de $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ par le nombre 1.

Alors,

Proposition 4.1.5.1

$\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est faiblement mais non fortement battable.

Démonstration : Soit $s(\cdot)$ une stratégie qui utilise uniquement les termes pairs de $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$. Selon la proposition 4.1.1.2, la différence de gains entre $s(\cdot)$ et la “buy and hold” est donnée par :

$$r_n/r_n^* = 2^{\sum_{i=1, i=2m+1}^n -\sigma_i} \quad m \in \mathbb{N} \quad (4.19)$$

Comme toute sous-suite raisonnable d’une suite de Martin-Löf reste une suite de Martin-Löf, les termes impairs de $\sigma_0, \sigma_1, \dots, \sigma_n$ constituent aussi une suite de 1 et -1 aléatoire dans le sens de Martin-Löf.

Selon la troisième propriété des suites de Martin-Löf rappelée à la page 190, la limite inf de la somme d’une suite de Martin-Löf⁸ est $-\infty$, nous avons donc

$$\liminf_{n \rightarrow \infty} \sum_{i=1, i=2m+1}^n \sigma_i = -\infty$$

Ainsi,

$$\limsup_{n \rightarrow \infty} r_n/r_n^* = \limsup_{n \rightarrow \infty} 2^{\sum_{i=1, i=2m+1}^n -\sigma_i} = +\infty$$

$s(\cdot)$ est donc une stratégie calculable qui peut faiblement battre la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n$.

Maintenant, nous montrons que $\sigma_0, \sigma_1, \dots, \sigma_n$ n’est fortement battable par aucune stratégie calculable. Cette démonstration utilise un raisonnement par l’absurde.

Soit $c_0, c_1, \dots, c_n, \dots$ le cours correspondant à la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$. Supposons qu’il existait une stratégie calculable $s(c_0, c_1, \dots, c_{n-1})$ qui peut la battre fortement, nous avons alors,

$$\liminf_{n \rightarrow \infty} r_n/r_n^* = 2^{\sum_{i=0, s(c_0, c_1, \dots, c_{i-1})=0}^n -\sigma_i} = +\infty \quad (4.20)$$

Où $s(c_0, c_1, \dots, c_{i-1}) = 0$ signifie que la rentabilité σ_i sera refusée.

À partir de $s(\cdot)$ nous pouvons construire une nouvelle stratégie $s'(\cdot)$ qui, parmi les termes refusés par $s(\cdot)$, refuse uniquement les termes impairs de $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, c’est-à-dire que :

$$s'(c_0, c_1, \dots, c_{i-1}) = \begin{cases} 0 & \text{si } i = 2m + 1 \text{ et } s(c_0, c_1, \dots, c_{i-1}) = 0 \\ 1 & \text{sinon} \end{cases}$$

Où $s'(c_0, c_1, \dots, c_{i-1}) = 0$ signifie que σ_i sera refusé par $s'(\cdot)$, et $s'(c_0, c_1, \dots, c_{i-1}) = 1$ le cas contraire.

La différence de gains entre $s(\cdot)$ et $s'(\cdot)$ se trouve uniquement dans les termes pairs de la suite $\sigma_0, \sigma_1, \dots, \sigma_n$. Comme les termes pairs de cette dernière suite

8. Ici, le terme “suite de Martin-Löf” nous renvoie à la notion des suites de Martin-Löf dont les 0 sont remplacés par les -1

sont toujours 1, $s'(\cdot)$ refuse moins de 1 que $s(\cdot)$. La richesse engendrée par $s'(\cdot)$ est donc supérieure ou égale à $s(\cdot)$.

Étant donné $s(\cdot)$ bat fortement $\sigma_0, \sigma_1, \dots, \sigma_n$, $s'(\cdot)$ bat fortement cette suite aussi.

Ainsi, la somme des rentabilités refusées par $s'(\cdot)$ doit vérifier :

$$\liminf_{n \rightarrow \infty} \sum_{i=1, s'(c_0, c_1, \dots, c_{i-1})=0}^n -\sigma_i = \infty \quad (4.21)$$

C'est-à-dire que

$$\limsup_{n \rightarrow \infty} \sum_{i=1, s'(c_0, c_1, \dots, c_{i-1})=0}^n \sigma_i = -\infty \quad (4.22)$$

Puisque les rentabilités refusées par $s'(\cdot)$ forment une sous-suite raisonnable de Martin-Löf, l'équation (4.22) est en contradiction avec la troisième propriété de la suite de Martin-Löf (1966) rappelée à la page 190. Il ne peut donc pas exister de stratégie calculable qui peut fortement battre $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$. ■

4.1.6 Liens entre des variables *i.i.d.* et les cours imbattables

Comme nous l'avons évoqué dans la section 4.1.4 (page 190), il existe une relation mathématique entre les suites de Martin-Löf (1966) et les suites réalisées par une collection de variables aléatoires *i.i.d.* $U(0, 1)$, dans la mesure où cette dernière variable engendre des suites de Martin-Löf avec une probabilité de 100%.

Fondée sur cette relation, nous pouvons démontrer la proposition suivante :

Soit $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ une suite de rentabilités réalisée par une collection de variables aléatoires *i.i.d.* $U(0, 1)$, et $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ la suite des rentabilités obtenue en remplaçant tous les termes pairs de $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ par le nombre 1.

Proposition 4.1.6.1

Avec une probabilité de 100%, $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ était faiblement mais non fortement battable par les stratégies calculables.

Démonstration : Selon la proposition 4.1.5.1 (page 195), si $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ est une suite de Martin-Löf, alors la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ obtenue en remplaçant tous les termes pairs de $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ par le nombre 1, est une suite faiblement mais non fortement battable. Comme une collection de variables aléatoires *i.i.d.* $U(0, 1)$ engendre des suites de Martin-Löf avec une

4.1. Modèle “tout acheter et tout vendre”

probabilité de 100%, c’est avec cette même probabilité que $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est faiblement mais non fortement battable par les stratégies calculables. ■

En finance, les rentabilités financières sont souvent supposées suivre une loi normale. Il est donc intéressant pour nous d’établir un lien entre les suites de rentabilités *i.i.d.* $N(0, \theta)$ et les suites imbattables au sens de la définition 4.1.0.4.

Or, ce lien entre la théorie des probabilités et la théorie de la calculabilité n’est pas facile à obtenir, car les suites engendrées par des collections de variables aléatoires *i.i.d.* $N(0, \theta)$ ne correspondent qu’exceptionnellement à des séries de cours entiers.

Pour résoudre ce problème, deux solutions peuvent être envisagées :

1. Soit nous modifions la définition de $c_0, c_1, \dots, c_n, \dots$ en imposant $c_n \in \mathbb{R}$ au lieu de $c_n \in \mathbb{N}^*$. Cette solution est peu satisfaisante, car les cours boursiers réels ont un niveau de précision fini.

Sur la plupart des places boursières, l’unité de cotation (“*tick size*”) est fixée à un centime ou au millième, en multipliant les cours observés par 100 (resp. par 1000), on obtient des nombres entiers. Par exemple, le cours 54,85 (euros) peut être considéré comme un nombre entier 5485. Les prix des actifs tels qu’on observe sur les marchés financiers sont donc mieux pris en compte par les nombres entiers que les nombres réels.

2. Soit nous modifions la relation entre une série de cours et sa suite de rentabilités :

Étant donné c_{n-1} , c_n est calculé selon l’équation 4.23.

$$c_n = \text{arrond}(c_{n-1}2^{\sigma_n}) \quad (4.23)$$

Où σ_n est tiré d’une loi normale $N(\mu, \theta)$. μ et θ sont des paramètres données. Et $\text{arrond}(c_{n-1}2^{\sigma_n})$ une fonction qui arrondit la valeur de $c_{n-1}2^{\sigma_n}$ au nombre entier le plus proche.

Ce faisant, même si les nombres réels tirés de $N(\mu, \theta)$ ne garantissent pas l’appartenance de c_n à \mathbb{N}^* , la fonction $\text{arrond}(\cdot)$ permet d’introduire la loi normale dans des simulations de séries boursières.

Remarquons que cette étape supplémentaire fait disparaître la correspondance stricte entre une série de cours et sa suite de rentabilités. Avec la fonction $\text{arrond}(\cdot)$, chaque suite de rentabilités correspond bien à une seule série de cours, en revanche, une série de cours peut correspondre à un nombre **infini** de suites de rentabilités.

Voici un exemple pour illustrer la relation “rentabilités-cours” caractérisée par l’équation 4.23.

Soit $\sigma_1 = 0,99$, $\sigma_2 = -1,01$, $\sigma_3 = 0,23$, trois rentabilités réalisées par une collection de variables aléatoires $i.i.d.N(0, 1)$. En imposant $\sigma_0 = c_0 = 10$, on peut transformer les rentabilités en séries de cours comme suit :

$$\begin{aligned} c_1 &= \text{arrond}(c_0 2^{0,99}) = 20 \\ c_2 &= \text{arrond}(c_1 2^{-1,01}) = 10 \\ c_3 &= \text{arrond}(c_2 2^{0,23}) = 12 \end{aligned}$$

À l'aide de cette nouvelle correspondance, on peut établir une relation entre une collection de variables aléatoires $i.i.d.N(0, \theta)$ et les suites imbattables.

Soit :

- $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ une suite de nombres réels réalisée par une collection de variables aléatoires $i.i.d.N(0, 1)$.
- $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ la suite des rentabilités obtenue en remplaçant les termes pairs de $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ par 1.
- $c_0, c_1, \dots, c_n, \dots$ une série de cours obtenue selon le principe :

$$c_0 = \sigma_0, \quad c_n = \text{arrond}(c_{n-1} 2^{\sigma_n})$$

Proposition 4.1.6.2

$c_0, c_1, \dots, c_n, \dots$ est faiblement mais non fortement battable avec une probabilité de 100%.

Démonstration : D'abord, les termes pairs de la suite $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ sont substitués par 1 pour faire en sorte que $c_n \in \mathbb{N}^*$.

Pour montrer que $c_0, c_1, \dots, c_n, \dots$ est faiblement mais non fortement battable avec une probabilité de 100%, il suffit de prouver :

1. $c_0, c_1, \dots, c_n, \dots$ est faiblement battable par la stratégie "tenir des titres aux termes pairs et du “cash” aux termes impairs" avec une probabilité de 100%.

Cette proposition est valable car la différence de gains entre une telle stratégie et la “buy and hold” est donnée par :

$$r_n / r_n^* = \sum_{i=1, i=2m+1}^n -\sigma_i, \quad m \in \mathbb{N} \tag{4.24}$$

Comme les termes impairs de la suite des rentabilités sont engendrés par une collection de variables aléatoires $i.i.d.N(0, 1)$, leur somme doit vérifier :

$$\limsup_{m \rightarrow \infty} \sum_{i=1, i=2m+1}^n -\sigma_i = \infty \tag{4.25}$$

La stratégie qui n'utilise que les termes pairs de $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ bat faiblement cette suite.

2. Est nulle la probabilité d’extraire raisonnablement une sous suite de $\sigma_1, \sigma_3, \sigma_5, \dots, \sigma_{2m+1}, \dots$ (notée σ_i^s) satisfaisant :

$$\liminf_{n \rightarrow \infty} \sum_{i=1}^n -\sigma_i^s = \infty \quad (4.26)$$

Comme la suite $\sigma_1, \sigma_3, \sigma_5, \dots, \sigma_{2m+1}, \dots$ est engendrée par une collection de variables aléatoires *i.i.d.N*(0, θ), c’est avec une probabilité de 100% que ses sous suites raisonnables satisfont,

$$\limsup(\text{resp. inf})_{n \rightarrow \infty} \sum_{i=1}^n \sigma_i^s = \infty(\text{resp. } -\infty) \quad (4.27)$$

Étant donnée la contradiction entre les équations (4.27) et (4.26), la dernière équation n’a nulle chance d’être vérifiée. ■

Ici, remarquons que la fonction *arrond*(.) est introduite pour étudier les suites *i.i.d.N*. Afin de rétablir la correspondance stricte entre une suite de rentabilités et sa série de cours, nous revenons au principe de transformation $\sigma_0 = c_0$ et $\sigma_n = \log_2(c_n) - \log_2(c_{n-1})$ pour tout $n \in \mathbb{N}^*$.

Les suites de Martin-Löf et des collections de variables aléatoires *i.i.d.N*(0, 1) sont des références classiques des suites aléatoires. Or, nous avons montré dans la proposition précédente que certaines suites fondées sur ces deux séries aléatoires sont faiblement battables par des stratégies calculables.

Jusqu’à présent, les seules suites fortement imbattables que nous avons présentées sont les séries strictement croissantes qui ne satisfont pas la condition 4.15 (page 188). Il est donc naturel de se demander ici, s’il existe des séries de cours infinies satisfaisant la condition 4.15 qui sont fortement imbattables ? La proposition suivante donne une réponse positive à cette question.

Soit $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, une suite de rentabilités obtenue en remplaçant les “0” par -1 et les “1” par 2 d’une suite de Martin-Löf, alors,

Proposition 4.1.6.3

$\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est une suite fortement imbattable qui satisfait la condition 4.15.

Démonstration : $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ satisfait la condition 4.15, car il existe un nombre infini de 0 dans une suite de Martin-Löf, ce qui implique un nombre infini de -1 dans $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$.

Maintenant, nous démontrons que $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est fortement imbattable avec un raisonnement par l’absurde.

4.1. Modèle “tout acheter et tout vendre”

Supposons que $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ soit faiblement battable. Selon la proposition 4.1.1.2, il existe une sous-suite raisonnable de $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ qui vérifie

$$\limsup_{n \rightarrow \infty} \sum_{i=1}^n -\sigma_i^s = \infty \quad (4.28)$$

Or, toute sous-suite raisonnable d’une suite de Martin-Löf reste une suite de Martin-Löf. Cela implique que $\sigma_1^s, \sigma_2^s, \dots, \sigma_n^s$ doit être une suite composée de 2 et -1 aléatoire dans le sens de Martin-Löf. Selon la deuxième propriété des suites de Martin-Löf rappelée à la page 190, la fréquence limite de -1 (resp. 2) est de 0,5. Cela entre en contradiction avec l’équation (4.28). $\sigma_0, \sigma_1, \dots, \sigma_n$ ne peut donc pas être faiblement battable. ■

Comme une collection de variables aléatoires $i.i.d.U(0,1)$ produit des suites de Martin-Löf avec une probabilité de 100%, on obtient la proposition 4.1.6.4.

Soit $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, une suite de rentabilités engendrée par une collection de variables aléatoires $i.i.d.U(-1,2)$, alors,

Proposition 4.1.6.4

$\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est fortement imbattable avec une probabilité de 100%.

De même, nous pouvons démontrer la proposition 4.1.6.1, en adoptant le principe de transformation $c_0 = \sigma_0$ et $c_n = \text{arrond}(c_{n-1}2^{\sigma_n}), n \in \mathbb{N}^*$.

Proposition 4.1.6.1

Soit $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, une suite de nombres réels engendrée par une collection des variables aléatoires $i.i.d.N(\mu, \theta)$, où μ, θ sont deux paramètres positifs, alors, la série de cours $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est fortement imbattable avec une probabilité de 100%.

Démonstration : Avec une probabilité de 100%, la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ satisfait :

$$\lim_{n \rightarrow \infty} \sum_{i=0}^n \sigma_i = \infty \quad (4.29)$$

Comme la série $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est engendrée de manière indépendante, toutes ses sous-suites raisonnables doivent vérifier cette même propriété avec une probabilité de 100%.

Étant donné que toute suite correspondant à un cours faiblement battable doit disposer d’une sous-suite raisonnable $(\sigma_0^s, \sigma_1^s, \dots, \sigma_n^s, \dots)$ satisfaisant :

$$\limsup_{n \rightarrow \infty} \sum_{i=0}^n -\sigma_i^s = \infty \quad (4.30)$$

Etant donnée la contradiction entre les équations (4.29) et (4.30), la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est fortement imbattable avec une probabilité de 100%. ■

Cette proposition est particulièrement intéressante pour les chercheurs en finance, car les séries de rentabilités sont souvent modélisées par l'équation suivante :

$$\log(c_n) = \log(c_{n-1}) + \sigma_n \quad (4.31)$$

où σ_n est une collection de variables aléatoires *i.i.d.* $N(\mu, \theta)$ avec $\mu, \theta > 0$.

Ici, μ est supposé positif car d'une manière générale, les cours bousiers ont une tendance haussière sur le long terme, et cela grâce aux rentabilités exigées par les investisseurs. La proposition 4.1.6.1 rejoint donc nos analyses du premier chapitre : si les rentabilités d'équilibre sont positives, et les rentabilités excessives sont imprévisibles, alors aucune stratégie de “*market timing*” ne doit battre la “*buy and hold*”.

4.2 Politique du blocage à 1 centime d'euro

À présent, aucune contrainte n'est posée sur la continuité des stratégies d'investissement. Or, en pratique, une stratégie est toujours soumise à une contrainte supplémentaire : un capital minimal est exigé à chaque fois qu'une stratégie veut “entrer sur le marché”. Plus précisément, afin d'exécuter un ordre d'achat à l'instant n , $d(\cdot)$ doit avoir assez de “*cash*” pour acquérir le titre coté au prix c_n .

Dans le modèle initial (page 182), cette contrainte n'est pas mentionnée, car nous avons supposé une divisibilité infinie des titres financiers : si à l'instant n , $d(\cdot)$ ne dispose pas assez de liquidité pour acheter une action en entier, il lui est possible d'en acheter $\frac{1}{2}$, $\frac{1}{10}$ ou $\frac{1}{1000}$.

Cette hypothèse est défendable, car les courtiers peuvent regrouper plusieurs commandes en un seul ordre d'achat à exécuter. Ce principe est inscrit dans la logique même du système de l'actionnariat. Tout individu n'a pas accès à la propriété entière des entreprises dont le droit d'entrée est souvent très élevé. Or, grâce au système de l'actionnariat, nous pouvons diviser la propriété d'une entreprise en petites parties (les actions) pour que chacun puisse les acquérir selon ses moyens.

Comme la richesse engendrée par une stratégie donnée reste toujours positive⁹, $d(\cdot)$ peut rentrer sur le marché à tout instant, car celle-ci peut n'acheter qu'une fraction infime d'un titre coté.

En réalité, il faut reconnaître que les actifs financiers ne peuvent pas être divisés en autant de parties que nous souhaitons. Par exemple, il est absurde d'investir moins de "1" centime d'euro sur le marché Euronext.

Face à une série de cours, si, à un instant donné, une stratégie $d(\cdot)$ se trouve avec une liquidité inférieure à 1 centime d'euro, alors elle ne pourra plus jamais poursuivre son activité d'investissement.

Pour prendre en compte cette contrainte réelle, nous allons introduire une politique du blocage à 1 centime d'euro¹⁰ dans notre modèle de stratégie, ce qui rendra finie la durée de vie de certaines stratégies sur une série de cours donnée.

Politique de blocage 4.2.0.1

Soit l_n la liquidité engendrée par la stratégie $d(\cdot)$ à l'instant n , sur le cours $c_1, c_2, \dots, c_n, \dots$. d_n est dite "bloquée" et son activité d'investissement s'arrête à jamais ssi $0 < l_n < 0,01$.

Dans le cadre du modèle "tout acheter et tout vendre", la richesse résultant d'une stratégie est soit 100% en "cash", soit 100% en titres. La condition $0 < l_n < 0,01$ fixée par la politique de blocage a donc deux implications :

1. à l'instant n , la richesse du portefeuille est en "cash",
2. la valeur de cette liquidité est inférieure à 1 centime euro.

Avant d'introduire la politique de blocage, face à une série de cours donnée, les stratégies calculables peuvent être divisées en deux catégories :

1. celles qui peuvent au moins faiblement battre la série de cours.
2. et celles qui ne le peuvent pas.

Après son introduction, chacune de ces deux catégories est divisée en deux groupes, ce qui donne 4 groupes au total :

1. certaines stratégies battant la série donnée sont bloquées.
2. certaines stratégies battant la série donnée continuent infiniment.
3. certaines stratégies ne battant pas la série donnée sont bloquées.
4. certaines stratégies ne battant pas la série donnée continuent infiniment.

9. Comme $r_n = 2^{\sum_{i=0}^n -\sigma_i^u}$, où σ_i^u représente toutes les rentabilités utilisées par la stratégie, quelle que soit la valeur de σ_i^u , r_n reste positive.

10. Ici, l'hauteur du seuil a été choisi de façon arbitraire. Nous allons montrer dans la proposition 4.2.0.4 que ce choix a relativement peu d'impact sur la possibilité de battre une série donnée.

En ce qui nous concerne, les stratégies de la première catégorie risquent de modifier la nature (battable ou pas) des séries de cours. Leur existence sera donc le sujet de la prochaine proposition.

Soit :

- $a_0, a_1, \dots, a_n, \dots$, une suite obtenue en remplaçant les 0 par les -1 dans une suite de Martin-Löf.
- $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, une suite de rentabilités obtenue en remplaçant les a_{2m+1} par les $-a_{2m}$ ($m \in \mathbb{N}$).
- $d(\cdot)$, une stratégie qui utilise les termes impairs de la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$.

Proposition 4.2.0.2

$d(\cdot)$ bat faiblement $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ sans politique du blocage à 1 centime d'euro, et sera bloquée après introduction de celle-ci.

Démonstration : Selon la proposition 4.1.1.2, la différence de gains entre $d(\cdot)$ et la “*buy and hold*” est calculée comme suit :

$$r_n / r_n^* = 2^{\sum_{i=1, i=2m}^n -\sigma_i} \tag{4.32}$$

Par construction, les termes pairs de la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ constituent une suite de -1 et de 1 aléatoire dans le sens de Martin-Löf. Ils satisfont donc :

$$\limsup_{n \rightarrow \infty} \sum_{i=1, i=2m}^n \sigma_i = \infty \tag{4.33}$$

$\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est donc faiblement battue par $d(\cdot)$ sans la politique de blocage.

Nous démontrons maintenant que $d(\cdot)$ sera bloquée après introduction de la politique du blocage à 0,01 euro.

Les rentabilités utilisées par $d(\cdot)$ sont les termes impairs de $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$. Comme $\sigma_{2m+1} = -\sigma_{2m}$, ces termes impairs constituent également une suite de -1 et de 1 , aléatoire dans le sens de Martin-Löf. La richesse accumulée $r_n = 2^{\sum_{i=1, i=2m+1}^n \sigma_i}$ vérifie donc :

$$\liminf_{n \rightarrow \infty} r_n = \liminf_{n \rightarrow \infty} 2^{\sum_{i=1, i=2m+1}^n \sigma_i} = 0$$

Selon la définition de la limite inf, $\forall i > 0, \exists M > i$ satisfaisant $r_M < 0,01$. De plus, comme r_n est en liquide à la fin de chaque instant impair, la stratégie $d(\cdot)$ est bloquée par la politique du blocage à 1 centime d'euro. ■

Remarquons que dans cet exemple, le blocage de $d(\cdot)$ n'influence pas la nature battable de la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, il existe des stratégies non bloquées qui peuvent battre cette suite de rentabilités, comme nous allons le démontrer dans la proposition suivante.

Soit

- $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$, la suite des rentabilités définie comme dans la proposition précédente.
- $s(\cdot)$, une stratégie qui refuse uniquement les termes impairs **précédés par un terme pair positif**.

Proposition 4.2.0.3

$s(\cdot)$ bat fortement $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ sans être bloquée par le seuil à 0,01 euro.

Démonstration : La différence de gains engendrée par $s(\cdot)$ est calculée par :

$$r_n/r_n^* = 2^{\sum_{i=1, i=2m+1, \sigma_{2m}=1}^n -\sigma_i}$$

Par construction, $\sigma_{2m+1} = -\sigma_{2m}$, les termes refusés par $s(\cdot)$ sont donc uniquement des -1 . On obtient ainsi

$$\lim_{n \rightarrow \infty} r_n/r_n^* = \lim_{n \rightarrow \infty} 2^{\sum_{i=1, i=2m+1, \sigma_{2m}=1}^n -\sigma_i} = \infty$$

De plus, quand n est suffisamment grand, la richesse accumulée de $s(\cdot)$ ne peut pas être inférieure à la richesse initiale, car les rentabilités négatives parmi les termes pairs de $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ sont toujours récompensées par les termes impairs qui les suivent. De plus, les termes pairs positifs font régulièrement augmenter r_n . La figure 4.1 est une comparaison graphique entre le cours c_n et la richesse accumulée r_n . ■

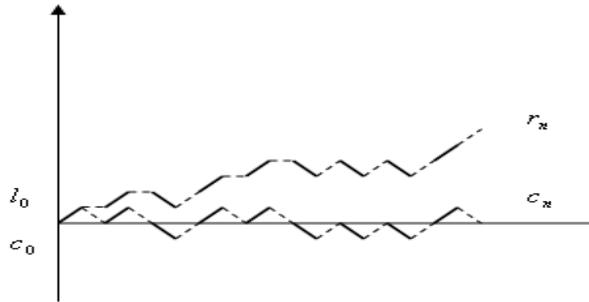


Fig. 4.1: c_n comparé avec la richesse engendrée r_n . Ici, sans perdre de généralité, nous imposons que $r_0 = c_0$. Nous constatons que la seule différence entre r_n et c_n se trouve dans les évolutions $\sigma_{2m+1} = -1$ représentées par les lignes discontinues.

Comme nous l'avons précisé au début de cette section, le seuil de blocage à 0,01 euro est un choix arbitraire. Nous montrons maintenant que la nature d'une série de cours (battable ou non) est indépendante de la valeur du seuil imposé.

Soit :

- $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ une suite de rentabilités donnée et $c_0, c_1, \dots, c_n, \dots$ la série de cours lui correspondant.
- $b_1 > b_2 \in \mathbb{R}^+$, deux seuils de blocage positifs.

Proposition 4.2.0.4

Si $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est faiblement battable en présence de la politique du blocage à b_2 , alors elle est aussi faiblement battable au seuil du blocage à b_1 .

Démonstration : Puisque $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est faiblement battable au seuil b_2 , il existe une stratégie calculable $d(\cdot)$ qui satisfait les deux conditions suivantes :

1. la liquidité engendrée par $d(\cdot)$, notée l_n , vérifie :

$$\liminf_{n \rightarrow \infty, l_n \neq 0} l_n = a, \quad a > b_2.$$

2. la différence de gains entre $d(\cdot)$ et la “*buy and hold*” satisfait :

$$\limsup_{n \rightarrow \infty} r_n / r_n^* = \infty$$

Selon la relation entre a et le seuil du blocage b_1 , nous distinguons deux cas de figure :

- $a > b_1$, selon la définition de la limite inf, $\exists M > 0$ pour que $d(\cdot)$ ne soit jamais bloquée au seuil b_1 après l’instant M . Si $d(\cdot)$ n’est pas bloquée non plus avant M , alors $d(\cdot)$ est une stratégie qui bat faiblement $c_0, c_1, \dots, c_n, \dots$ en présence de la politique du blocage à b_1 .

Si $d(\cdot)$ est bloquée avant M , alors en modifiant un nombre fini de fois les décisions de $d(\cdot)$, nous pouvons construire une stratégie calculable $d'(\cdot)$ qui peut faiblement battre $c_0, c_1, \dots, c_n, \dots$ sans être bloquée par le seuil à b_1 .

- $0 < a \leq b_1$.

Soit *begin*(\cdot) une stratégie optimale qui utilise toutes les augmentations et qui refuse toutes les baisses durant les n premiers termes de la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$. Comme $\sum_{i=1, \sigma_i < 0}^n \sigma_i = -\infty$ et $c_n \in \mathbb{N}^*$, la richesse engendrée par *begin*(\cdot), r'_n tend vers ∞ quand n augmente.

Soit $[l_0], [l_1], \dots, [l_n], \dots$ la sous-suite de $l_0, l_1, \dots, l_n, \dots$ qui contient les $l_i > 0$. Selon la définition de limite inf :

1. $\forall \varepsilon > 0, \forall m \in \mathbb{N}, \exists M > m, [l_M] < a + \varepsilon$
2. $\forall \varepsilon > 0, \exists m \in \mathbb{N}, \forall M > m, [l_M] > a - \varepsilon$

Il existe donc un instant M satisfaisant :

1. $r'_M > b_1$
2. la richesse engendrée par la $d(\cdot)$ à l’instant M satisfait : $r_M < a + \varepsilon$
3. $\forall n > M, [l_n] > a - \varepsilon$

La stratégie $d'(\cdot)$ qui prend les mêmes décisions que *begin*(\cdot) avant l’instant M , et qui suit $d(\cdot)$ après M , engendre une richesse de :

$$r'_n = r_n \frac{r'_M}{r_M} > \frac{(a - \varepsilon)b_1}{a + \varepsilon} \approx b_1$$

pour tout $n > M$.

$d'(\cdot)$ ainsi construite n'est pas bloquée au seuil b_1 face à la série $c_0, c_1, \dots, c_n, \dots$.

Étant donné que $r'_n \geq r_n$ pour tout $n \in \mathbb{N}^*$ et que $d(\cdot)$ bat faiblement $c_0, c_1, \dots, c_n, \dots$, la stratégie $d'(\cdot)$ doit aussi faiblement battre la série de cours $c_0, c_1, \dots, c_n, \dots$. ■

La proposition 4.2.0.4 montre l'indépendance entre la valeur du seuil de blocage imposée et la possibilité de battre une série donnée. Certes, en imposant un seuil élevé, nous "bloquons" plus de stratégies qu'avec un seuil bas. Cependant, si une série est faiblement battable à un seuil bas, alors qu'elle l'est aussi avec un seuil plus élevé.

4.3 Politique de blocage et nature imbattable des cours infinis

Une série de cours fortement (faiblement) battable en absence de politique de blocage resterait-elle fortement (faiblement) battable après son introduction ?

Avant de nous concentrer sur cette question, on montre d'abord une proposition préparatoire qui sera souvent utilisée dans les futures démonstrations.

Soit $c_1, c_2, \dots, c_n, \dots$ une série de cours fortement imbattable en présence de la politique de blocage. De plus, $c_1, c_2, \dots, c_n, \dots$ satisfait la condition 4.15 (page 188). Alors,

Proposition 4.3.0.1

$c_1, c_2, \dots, c_n, \dots$ satisfait :

$$\liminf_{n \rightarrow \infty} c_n = +\infty \quad (4.34)$$

Démonstration : Comme $c_n \in \mathbb{N}^*$, la limite inf de c_n , $a > 0$.

$\liminf_{n \rightarrow \infty} c_n = a$ implique que

1. $\forall \varepsilon > 0, \forall m \in \mathbb{N}, \exists M > m, c_M < a + \varepsilon$
2. $\forall \varepsilon > 0, \exists m \in \mathbb{N}, \forall M > m, c_M > a - \varepsilon$

Comme $c_n \in \mathbb{N}^*$, $\varepsilon = 0$.

On peut construire une stratégie calculable $d(\cdot)$ qui achète à chaque instant i satisfaisant : $c_i = a$ ($i \geq M$), et qui revend au premier instant j satisfaisant $c_j > a$ ($j > i$).

4.3. Politique de blocage et nature imbattable des cours infinis

Comme $c_n \geq a$ pour tout $n > M$, $d(\cdot)$ gagne à chacune de ses interventions, elle n'est donc jamais ruinée.

De plus, étant donné $\liminf_{n \rightarrow \infty} c_n = a$, $d(\cdot)$ doit intervenir un nombre infini de fois sur le marché. Par conséquent, quand $n \rightarrow \infty$, la richesse engendrée par $d(\cdot)$ tend vers $+\infty$.

D'un autre côté, les gains de la stratégie "buy and hold" (c_n) ont une limite inf $a < \infty$, la différence de gains entre les deux stratégies r_n/r_n^* doit donc satisfaire : $\limsup r_n/r_n^* = \infty$.

$d(\cdot)$ bat faiblement la série de cours $c_1, c_2, \dots, c_n, \dots$ en présence de la politique de blocage. ■

En effet, la proposition précédente illustre le principe suivant : en connaissant le prix minimal d'un titre donné ($\liminf_{n \rightarrow \infty} c_n < \infty$), on peut faiblement battre la "buy and hold" en achetant le titre à chaque fois son cours est proche du seuil minimal.

Pendant des périodes de crises, certains titres financiers valent entre 1 et 2 euros. Par exemple, cela fut le cas de l'action Dexia en 2008. En supposant la continuité d'exploitation de Dexia, on peut acheter cette action à 1 euro sans prendre beaucoup de risque. Bien sûr, en réalité, le risque de faillite est toujours associé à ce type de titres à bas prix. Le vrai arbitrage n'existe pas sur les marchés financiers.

Proposition 4.3.0.2

Une série de cours **fortement** battable en absence de la politique du blocage à 0,01 euro l'est également après son introduction.

Démonstration : Soit $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ une série de cours fortement battue par la stratégie $d(\cdot)$ en absence de politique de blocage. La différence de gains entre $d(\cdot)$ et la "buy and hold" doit satisfaire

$$\liminf_{n \rightarrow \infty} r_n/r_n^* = \infty \quad (4.35)$$

Comme $r_n^* = l_0 \frac{c_n}{c_0} > 0$, l'équation (4.35) implique :

$$\lim_{n \rightarrow \infty} r_n = \infty \quad (4.36)$$

Il existe donc un instant M qui satisfait $r_n > 0,01$ pour tout $n \geq M$. Autrement dit, si $d(\cdot)$ est bloquée par le seuil de 1 centime d'euro, alors elle l'est avant l'instant M .

Comme nous l'avons montré précédemment, pour les premiers M termes d'une suite donnée, nous pouvons toujours construire un début de stratégie *begin*(\cdot) qui utilise toutes les augmentations et qui refuse toutes les baisses.

La stratégie $d'(\cdot)$ qui combine $begin(\cdot)$ et $d(\cdot)$ doit engendrer une richesse accumulée supérieure à celui de $d(\cdot)$ à chaque instant. Comme $d'(\cdot)$ ne subit aucune perte avant l'instant M , elle ne peut pas être bloquée avant M .

De plus, comme $d(\cdot)$ n'est pas bloquée après l'instant M et la richesse engendrée par $d'(\cdot)$ est toujours supérieure à $d(\cdot)$, alors, $d'(\cdot)$ ne peut pas être bloquée après M non plus.

$d'(\cdot)$ est donc une stratégie qui bat fortement la suite $\sigma_1, \sigma_2, \dots, \sigma_n, \dots$ en présence de la politique du blocage à 1 centime d'euro. ■

Soit :

– $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ une suite de rentabilités dont la moyenne limite satisfait :

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \sigma_i}{n} = a \in \mathbb{R}$$

– $d(\cdot)$ une stratégie qui bat **faiblement mais non fortement** la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ en absence de politique de blocage.

Proposition 4.3.0.3

Si la sous-suite utilisée par $d(\cdot)$ notée $\sigma_0^s, \sigma_1^s, \dots, \sigma_n^s, \dots$ satisfait :

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \sigma_i^s}{n} = b \in \mathbb{R}$$

alors $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est aussi faiblement battable après l'introduction de la politique du blocage à $0 < bloc < \infty$ euros.

Démonstration : Soit $c_0, c_1, \dots, c_n, \dots$ la suite de cours correspondant à $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$. Selon la proposition 4.3.0.1, si $c_0, c_1, \dots, c_n, \dots$ n'est pas fortement battable par $d(\cdot)$, alors,

$$\liminf_{n \rightarrow \infty} c_n = \liminf_{n \rightarrow \infty} c_0 2^{\sum_{i=1}^n \sigma_i} = +\infty.$$

Par conséquent, la moyenne limite $a > 0$.

Soit $\overline{\sigma}_0^s, \overline{\sigma}_1^s, \dots, \overline{\sigma}_n^s, \dots$ (resp. $\sigma_0^s, \sigma_1^s, \dots, \sigma_n^s, \dots$) la sous-suite refusée (resp. utilisée) par $d(\cdot)$, nous avons donc

$$\{\overline{\sigma}_0^s, \overline{\sigma}_1^s, \dots, \overline{\sigma}_n^s, \dots\} \cup \{\sigma_0^s, \sigma_1^s, \dots, \sigma_n^s, \dots\} = \{\sigma_0, \sigma_1, \dots, \sigma_n, \dots\}.$$

$c_0, c_1, \dots, c_n, \dots$ est faiblement mais non fortement battue par $d(\cdot)$ implique que :

$$\liminf_{n \rightarrow \infty} \sum_{i=1}^n \sigma_i^s = -\infty. \tag{4.37}$$

Ainsi, la moyenne limite b , doit être inférieure ou égale à 0.

Étant donné $\{\overline{\sigma}_0^s, \overline{\sigma}_1^s, \dots, \overline{\sigma}_n^s\} \cup \{\sigma_0^s, \sigma_1^s, \dots, \sigma_n^s\} = \{\sigma_0, \sigma_1, \dots, \sigma_n\}$, $a > 0$ et $b \leq 0$, la moyenne limite de la sous-suite utilisée, $c = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \overline{\sigma}_i^s}{n} > 0$.

$d(\cdot)$ est donc une stratégie calculable dont la richesse accumulée r_n satisfait :

$$\lim_{n \rightarrow \infty} r_n = l_0 2^{\sum_{i=1}^n \overline{\sigma_i}} = \infty. \quad (4.38)$$

Par conséquent, quelle que soit la valeur du seuil $bloc < \infty$, il existe toujours un instant M qui satisfait $r_n > bloc$ pour tout $n > M$. Comme pour la démonstration précédente, en modifiant les premières M décisions à $d(\cdot)$, on peut construire une stratégie $d'(\cdot)$ dont la richesse accumulée $r'_n \geq r_n$ à chaque instant n ($d'(\cdot)$ bat donc faiblement la série $c_0, c_1, \dots, c_n, \dots$), et qui n'est pas bloquée par la politique du blocage à $bloc$ euros. ■

Remarquons que dans la section actuelle, nous n'avons pas pu apporter de réponse générale à la question "les suites faiblement battables sans la politique de blocage le sont-elles également après l'introduction de cette politique?" La proposition 4.3.0.3 ne traite cette question que dans un cadre très spécifique : la suite des rentabilités et la sous-suite utilisée doivent toutes deux avoir une moyenne limite convergente.

Certes, les suites des rentabilités satisfaisant cette condition ne représentent qu'une petite partie des suites de nombres réels, une généralisation de notre résultat est certainement intéressante sur le plan mathématique.

Cependant, d'un point de vue financier, la condition de convergence ne semble pas excessive : les premier et deuxième moments des séries de rentabilités sont souvent considérés convergents. La proposition 4.3.0.3 est donc une première réponse relativement satisfaisante à notre question du départ.

Sous les conditions de convergence, pour chaque stratégie calculable $d(\cdot)$ qui peut battre une série de rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ en absence de politique de blocage, il existe toujours un moment critique M à partir duquel $d(\cdot)$ ne sera plus jamais bloquée, et ce, quelle que soit la valeur du seuil introduite $bloc$. Ainsi, en modifiant les premières M décisions à $d(\cdot)$, on peut toujours construire une autre stratégie calculable $d'(\cdot)$ qui bat $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ sans être bloquée le seuil $bloc$.

Ce résultat semble fournir un support théorique aux "stop loss rules" pratiquées par certains fonds mutuels :

La règle du "stop loss" consiste à vendre un investissement donné dès que la perte accumulée atteint un certain seuil préfixé, et à le racheter après une remontée conséquente de l'actif. Cette pratique vise à limiter les scénarios catastrophes qu'on pourra parfois constater sur les marchés financiers.

Comme cette règle incite les investisseurs à “vendre aux prix faibles et à acheter aux prix forts”, son effet sur leur performance à long terme est souvent considéré négatif par les chercheurs en finance. Kaminski et Lo (2007) démontrent pourtant que “*stop loss rules*” peuvent améliorer la performance des fonds si les rentabilités successives sont positivement auto-corrélées.

Or, sans supposer une telle auto-corrélation, nous avons montré que sous certaines conditions de convergence, si une stratégie, $d(\cdot)$ bat une série de cours donnée, alors interrompre $d(\cdot)$ un nombre fini de fois ne l’empêchera pas de battre la série à long terme : $d(\cdot)$ finira par arriver à un moment critique, M à partir duquel son gain accumulé ne franchira plus jamais le seuil de blocage.

4.4 Généralisation du modèle “tout acheter et tout vendre”

Jusqu’à présent, à chaque stratégie considérée, nous imposons systématiquement la condition “tout acheter ou tout vendre”. Or, en réalité, les investisseurs varient souvent leur “mise” en fonction de leur état financier ou de leur niveau de confiance à un instant donné.

Afin de rendre plus réaliste notre modèle de stratégie, nous introduisons maintenant la possibilité de choisir la proportion de richesse engagée dans chaque intervention sur le marché.

- Soit $c_0, c_1, \dots, c_n, \dots$ ($c_n \in \mathbb{N}^*, n \in \mathbb{N}$) une suite de cours donnée.
- Soit $d(c_0, c_1, \dots, c_{n-1}) \in \mathbb{Q}$ (désormais d_n) la fonction de décision qui donne le ratio $\frac{\text{valeur des titres détenus}}{\text{valeur de la richesse totale}}$ **à la fin** de la période n .
- Soit $q(c_0, c_1, \dots, c_{n-1}) \in \mathbb{Q}$ (désormais q_n), le ratio $\frac{\text{valeur des titres détenus}}{\text{valeur de la richesse totale}}$ **au début** de l’instant n . Sans perdre de généralité, nous imposons $q_1 = 0$.

Ici, il convient de préciser que la différence entre q_n et d_n est que q_n ne prend pas en compte l’opération réalisée à l’instant n . q_n se distingue également de d_{n-1} , car de l’instant $n-1$ à n , le cours est passé de c_{n-2} à c_{n-1} , le ratio $\frac{\text{valeur des titres détenus}}{\text{valeur de la richesse totale}}$ est modifié suite à ce changement.

Par exemple,

1. supposons qu’à l’instant 1, $d_1 = 10\%$; nous allons utiliser 10% de la richesse initiale (l_0) pour acheter des actions au prix c_0 . À la fin de l’instant 1, la valeur des titres représente 10% de la valeur du portefeuille.
2. à l’instant 2, le prix de l’actif devient c_1 , la valeur relative des actifs détenus au début de l’instant 2 est donnée par :

$$q_2 = \frac{d_1 \frac{c_1}{c_0}}{d_1 \frac{c_1}{c_0} + (1 - d_1)}$$

4.4. Généralisation du modèle “tout acheter et tout vendre”

Puis, en fonction des cours historiques c_0, c_1 , d_2 doit prendre une décision concernant l'acte à réaliser à l'instant 2; supposons ici que $d_2 = 5\%$.

Pour exécuter cet ordre, il faut comparer d_2 et q_2 ,

- si $q_2 < 5\%$, il faudrait acheter des titres au prix c_1 pour que leur valeur représente 5% de celle du portefeuille.
- si $q_2 > 5\%$, il faudrait vendre des titres au prix c_1 pour que ceux-ci ne représentent que 5% du portefeuille à la fin de l'instant 2.
- si $q_2 = 5\%$, aucune opération ne serait effectuée.

d_n, q_n ainsi définies vérifient les relations suivantes :

- Au départ, $d(\cdot)$ commence avec un capital initial en “*cash*”. C'est-à-dire que $l_0 > 0, d_0 = q_0 = 0$
- À l'instant 1, c_0 est connu, d_1 prend une décision selon ce premier cours de la série :

$$\begin{aligned} q_1 &= 0 \\ d_1 &\in \mathbb{Q}, \quad 0 \leq d_1 \leq 1 \end{aligned} \tag{4.39}$$

- À l'instant 2,

$$\begin{aligned} q_2 &= \frac{d_1 \frac{c_1}{c_0}}{d_1 \frac{c_1}{c_0} + (1 - d_1)} \\ d_2 &\in \mathbb{Q}, \quad 0 \leq d_2 \leq 1 \end{aligned} \tag{4.40}$$

- À l'instant n ,

$$\begin{aligned} q_n &= \frac{d_{n-1} \frac{c_{n-1}}{c_{n-2}}}{d_{n-1} \frac{c_{n-1}}{c_{n-2}} + (1 - d_{n-1})} \\ d_n &\in \mathbb{Q}, \quad 0 \leq d_n \leq 1 \end{aligned} \tag{4.41}$$

Remarquons que q_n est calculable si d_n l'est également. La richesse accumulée à l'instant n est donc donnée par :

$$r_n = l_0 \prod_{i=1}^n (d_i 2^{\sigma_i} + 1 - d_i) \tag{4.42}$$

Le gain accumulé de la “*buy and hold*” est de :

$$r_n^* = l_0 2^{\sum_{i=1}^n \sigma_i} \tag{4.43}$$

On constate ici que les stratégies définies par ce modèle (désormais “**stratégies généralisées**”) ont plus de liberté dans leurs décisions d'investissement que celles présentées la section 4.1 (page 182, désormais “**stratégies simples**”).

Les stratégies généralisées qui sont sans doute plus proches de la réalité que les stratégies simples peuvent éventuellement battre plus de suites de rentabilités. Cette question est étudiée dans les développements suivants.

4.5 Stratégies simples vs. stratégies généralisées

Nous comparons maintenant les stratégies généralisées et simples en fonction de leur capacité à battre des séries de cours entiers.

Proposition 4.5.0.1

Si une série de cours est faiblement (resp. fortement) battable par les stratégies simples, alors elle l'est également par les stratégies généralisées.

Démonstration : Les stratégies simples font partie des stratégies généralisées. Donc, une suite de cours faiblement (fortement) battue par une stratégie simple l'est également par une stratégie généralisée. ■

Proposition 4.5.0.2

Il existe des cours fortement battables par les stratégies généralisées qui sont fortement imbattables par les stratégies simples.

En finance, les stratégies géométriques (Aucamp 1993, Ross 1999, Christensen 2005) calculent la meilleure proportion à investir dans un actif risqué en maximisant la moyenne géométrique des rendements futurs.

Par exemple, la moyenne géométrique, $E(\log(X))$ du jeu de “pile ou face” décrit dans la figure 4.2 est calculée par l'équation (4.44).

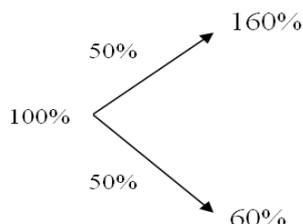


Fig. 4.2: Dans ce jeu de “pile ou face”, on a 50% de chance de faire un rendement de 60%, et 50% de chance de faire une perte de 40%.

$$E(\log(X)) = 50\% \log(160\%f + (1 - f)) + 50\% \log(60\%f + (1 - f)) \quad (4.44)$$

Où $E(\cdot)$ désigne la fonction d'espérance, f le pourcentage à investir dans le jeu. $E(\log(X))$ atteint son maximum avec $f^* = \frac{5}{12}$. La stratégie géométrique recommande donc à investir $\frac{5}{12}$ de son capital dans ce jeu de “pile ou face”.

On remarque ici que le pourcentage à investir dans une série dépend de la distribution des gains. Pour certaines séries de rentabilités, cette meilleure proportion n'est pas 100%, comme revendiquée par la "buy and hold". Dans ce cas de figure, les stratégies générales s'avèrent plus rentables que les stratégies simples. Nous allons donc utiliser les stratégies géométriques pour démontrer la proposition 4.5.0.2.

Démonstration : Soit $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ une suite composée de 2 et -1 aléatoire dans le sens de Martin-Löf. $c_0, c_1, \dots, c_n, \dots$ la série de cours correspondant à $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$. Nous montrons maintenant que $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est fortement battable par les stratégies généralisées mais **fortement imbattable** par les stratégies simples.

- Comme nous l'avons montré dans la proposition 4.1.6.3 (page 199), $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est fortement imbattable par les stratégies simples.
- Maintenant, nous montrons que $\sigma_1, \sigma_2, \dots, \sigma_n, \dots$ est fortement battable par des stratégies généralisées.

Soit f , la proportion de la richesse investie par une stratégie généralisée d_n , la moyenne géométrique des rendements de d_n est donnée par :

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \log \frac{r_n}{r_{n-1}}}{n} = 0,5 \log(2^2 f + 1 - f) + 0,5 \log(2^{-1} f + 1 - f) \quad (4.45)$$

Où \log désigne toujours le logarithme de base 2. Equation (4.45) atteint son maximum avec $f^* = \frac{5}{6}$.

En introduisant $f^* = \frac{5}{6}$, la stratégie d_n devient une stratégie géométrique. La moyenne de ses rendement sur $\sigma_1, \sigma_2, \dots, \sigma_n, \dots$ est calculée par :

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \log \frac{r_n}{r_{n-1}}}{n} = 0,5 \log\left(3 \times \frac{5}{6}\right) + 0,5 \log\left(1 - 0,5 \times \frac{5}{6}\right) = \log \frac{3,5^2}{6} > 1 \quad (4.46)$$

En même temps, le gain de la "buy and hold" est donné par :

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \log \frac{r_n^*}{r_{n-1}^*}}{n} = 1 \quad (4.47)$$

Étant donné que

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \log \frac{r_n}{r_{n-1}}}{n} > \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \log \frac{r_n^*}{r_{n-1}^*}}{n} \quad (4.48)$$

d_n bat fortement la suite de rentabilités $\sigma_1, \sigma_2, \dots, \sigma_n, \dots$. ■

4.6 La vitesse de convergence des stratégies d'investissement

Un euro d'aujourd'hui n'est pas égal à un euro de demain, tel est le principe économique qui domine la méthode d'actualisation. La vitesse à laquelle une stratégie engendre des bénéfices est un critère important pour l'évaluer.

Ceci a été noté par Grinold et Kahn (1999)¹¹, La notion “*d’information ratio*”, proposée dans ce manuel, qui mesure à la fois “*our ability to forecast each asset’s residual return*” et “*the number of times per year that we can use our skill* (Grinold et Kahn 1999, p.6)”, est devenu un des outils les plus répandus dans l’évaluation des fonds d’investissement.

Afin de compléter notre étude des cours imbattables, la vitesse à laquelle une stratégie d’investissement bat une série de cours fait le sujet de discussion dans cette section.

Dans un premier temps, la vitesse à laquelle une différence de gains¹² converge vers sa $\lim \sup / \lim \inf$, nommée “vitesse de convergence”, sera utilisée pour mesurer la rapidité d’une stratégie donnée.

Ensuite, nous allons étudier la relation théorique entre la rapidité d’une stratégie et sa performance finale. Nous montrons que la vitesse de convergence n’implique pas sa performance à long terme (faiblement ou fortement battre). Ces deux aspects d’une stratégie d’investissement sont indépendants.

Définition 4.6.0.1

Une série de cours $c_0, c_1, \dots, c_n, \dots$ est dite “battue par la stratégie $d(.)$ à la vitesse v ”, si la différence de gains entre $d(.)$ et la stratégie “buy and hold” vérifie :

$$\lim_{n \rightarrow \infty} \frac{\log(r_n/r_n^*)}{n^v} = a \tag{4.49}$$

Où “ a ” est une constante positive.

Remarquons que la vitesse de convergence ainsi définie indique la fréquence à laquelle $d(.)$ surperforme la “buy and hold”. Quand $v = 1$, l’équation (4.49) se transforme en équation (4.50)

$$\lim_{n \rightarrow \infty} \log(\sqrt[n]{r_n/r_n^*}) = a > 0 \tag{4.50}$$

Cela implique que la moyenne géométrique de la différence de gains r_n/r_n^* est supérieure à 1. Autrement dit, si une stratégie bat un cours à la vitesse 1, alors, elle doit engendrer, en moyenne, une différence de gains positive à chaque instant passé.

Quand $v = \frac{1}{2}$, l’équation (4.49) devient

$$\lim_{n \rightarrow \infty} \log(\sqrt[n]{\sqrt[n]{r_n/r_n^*}}) = a > 0 \tag{4.51}$$

11. Un des manuels de référence pour les professionnels d’investissement.

12. Entre une stratégie active et la “buy and hold”.

L'équation (4.51) signifie que r_n/r_n^* augmente en moyenne tous les $n^{\frac{1}{2}}$ termes.

Une vitesse $v > 1$ indique que non seulement la différence de gains augmente à chaque instant passé, de plus, l'ampleur de cette augmentation croît avec n .

Par conséquent, plus la valeur de v est importante, plus vite la stratégie $d(\cdot)$ bat la série de cours donnée.

En combinant cette notion de vitesse avec celle de cours faiblement (resp. fortement) battable, nous obtenons la définition d'un cours faiblement (resp. fortement) battable à la vitesse v .

Définition 4.6.0.2

Une série de cours $c_0, c_1, \dots, c_n, \dots$ est dite "faiblement (resp. fortement) battable à la vitesse v ", s'il existe une stratégie calculable $d(\cdot)$, qui la bat faiblement (resp. fortement) à la vitesse v .

Par exemple, la séries de cours $2, 1, 2, 1, 2, \dots, 2, 1, \dots$ qui commence par 2, et qui se multiplie (resp. divise) par 2 à chaque terme pair (resp. impair), est battue **à la vitesse** $v = 1$ par la stratégie $d(\cdot)$ qui "achète aux termes impairs et qui vend aux termes pairs". La différence de gains entre $d(\cdot)$ et la "buy and hold" satisfait :

$$\log(r_n/r_n^*) = \begin{cases} \frac{n}{2} & \text{si } n \text{ est un nombre pair} \\ \frac{n+1}{2} & \text{si } n \text{ est un nombre impair} \end{cases} \quad (4.52)$$

Comme $\lim_{n \rightarrow \infty} r_n/r_n^* = \infty$, la série de cours $2, 1, 2, 1, 2, \dots, 2, 1, \dots$ est fortement battue par $d(\cdot)$. En même temps, comme $\lim_{n \rightarrow \infty} \frac{\log(r_n/r_n^*)}{n^1} = \frac{1}{2} > 0$, $d(\cdot)$ bat la suite $2, 1, 2, 1, 2, \dots, 2, 1, \dots$ à la vitesse $v = 1$.

Soit :

- $a_0, a_1, a_2, \dots, a_n, \dots$ une suite composée de 1 et -1 aléatoire dans le sens de Martin-Löf.
- $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_n, \dots$ une suite de rentabilités obtenue en remplaçant par 1 les termes pairs de la suite $a_0, a_1, a_2, \dots, a_n, \dots$.

Proposition 4.6.0.3

La stratégie $d(\cdot)$, qui utilise les termes pairs de la suite $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_n, \dots$, la bat faiblement à la vitesse $\frac{1}{2}$.

Démonstration : Comme la différence de gains entre une stratégie donnée et la "buy and hold" dépend de la somme des rentabilités refusées, celle entre $d(\cdot)$ et la "buy and hold" est calculée par :

$$\log(r_n/r_n^*) = 2^{\sigma_1 + \sigma_3 + \dots + \sigma_n} \quad (4.53)$$

Comme les termes impairs de la série $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_n, \dots$ forment une suite de -1 et de 1 aléatoire dans le sens de Martin-Löf, la somme $\sum_{i=1, i=2m+1}^n \sigma_i$ ($m \in \mathbb{N}$) doit satisfaire :

1. la $\lim \sup$ (resp. $\lim \inf$) de la somme est ∞ (resp. $-\infty$), c'est-à-dire que

$$\limsup_{n \rightarrow \infty} (\text{resp. } \liminf_{n \rightarrow \infty}) \sum_{i=1, i=2m+1}^n \sigma_i = +\infty (\text{resp. } -\infty) \quad (4.54)$$

2. la somme s'écarte de 0 à la vitesse \sqrt{m} . C'est-à-dire que

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1, i=2m+1}^n \sigma_i}{\sqrt{\lfloor \frac{n}{2} \rfloor}} = a \quad (4.55)$$

Où $a \in \mathbb{R}$ est une constante positive, et la fonction $\lfloor \frac{n}{2} \rfloor$ calcule le plus grand entier qui ne dépasse pas $\frac{n}{2}$.

Par définition, $d(\cdot)$ bat faiblement la suite $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_n, \dots$ à la vitesse $\frac{1}{2}$. ■

4.6.1 La performance à long terme et la vitesse de convergence, deux aspects complémentaires d'une stratégie d'investissement

À ce stade d'analyse, une stratégie d'investissement peut être évaluée selon deux critères : sa performance à long terme et la vitesse à laquelle elle atteint cette performance.

Étant donnée une série de cours infinie, une stratégie qui la bat **fortement** est intuitivement considérée comme “meilleure” que celle qui ne la bat que **faiblement**. Alors qu'en adoptant le critère de la rapidité, plus vite une stratégie bat un cours donné, “meilleure” elle sera.

Or, sur le plan théorique, ces deux grilles d'évaluation posent un problème méthodologique : la performance finale et la vitesse de convergence, sont-elles interdépendantes ? Une stratégie qui bat fortement une série de cours donnée, serait-elle toujours plus rapide qu'une autre stratégie qui ne la bat que faiblement ? Autrement dit, face à une série de cours donnée, peut-on fixer une vitesse critique, notée V , pour que toutes les stratégies qui convergent plus (resp. moins) vite que V puissent fortement (resp. faiblement) la battre ?

Cette section est destinée à apporter une réponse à ces questions générales. Plus précisément, nous montrons avec des exemples concrets que les notions “faiblement battre” et “fortement battre” ne sont pas directement liées avec celle de la

4.6. La vitesse de convergence des stratégies d'investissement

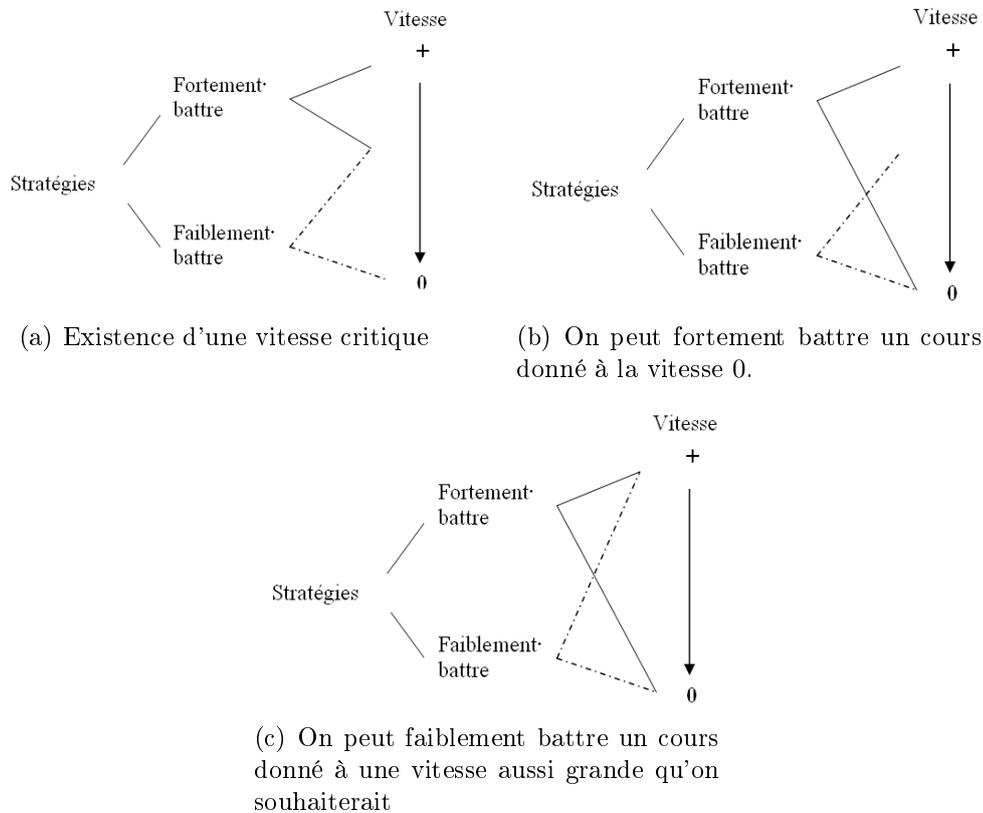


Fig. 4.3: Relation entre la performance finale et la vitesse de convergence.

vitesse de convergence. Une stratégie qui peut fortement battre une série de cours n'est pas toujours plus rapide qu'une stratégie qui ne la bat que faiblement. Les deux critères sont complémentaires, aucun des deux ne peut remplacer l'autre, car ils illustrent chacun un aspect différent des stratégies d'investissement.

La relation entre la performance d'une stratégie donnée et sa vitesse de convergence peut se trouver dans l'une des 3 catégories illustrées par la figure 4.3 (page 217) :

- Si la relation entre la performance et la vitesse est présentée par le schéma 4.3(a), alors les stratégies qui peuvent fortement battre une série de cours convergent toujours plus vite que celles qui ne la battent que faiblement. Il existe donc une vitesse critique, V , qui permet de déduire la performance finale d'une stratégie donnée selon sa vitesse de convergence.
- Si la relation entre la performance finale et la vitesse de convergence est donnée par le schéma 4.3(b), alors il est possible de fortement battre une série de cours à la vitesse 0. Cela implique l'absence de la vitesse critique V . En revanche, nous pourrions toujours fixer une vitesse V_2 , pour que toutes les stratégies convergeant plus vite que V_2 puissent fortement battre la série concernée. V_2 est différent de V , car il est possible de fortement battre un

cours à une vitesse moins élevée que V_2 .

- Si la relation entre la performance finale et la vitesse de convergence est caractérisée par le schéma 4.3(c), alors il n'existe aucune relation directe entre la vitesse et la performance d'une stratégie donnée. Non seulement il existe des cours fortement battables à la vitesse 0, certaines stratégies peuvent même faiblement battre un cours à une vitesse infinie. La performance et la vitesse sont deux aspects complémentaires d'une stratégie donnée.

Maintenant que nous avons évoqué ces différentes possibilités, commençons à identifier la "vraie relation" entre ces deux critères.

Proposition 4.6.1.1

Il existe des séries de cours fortement battables à la vitesse 0.

Démonstration : Soit $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$, une suite de Martin-Löf. Sans perdre de généralité, nous supposons que les premiers 7 termes de $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$ sont 1.

Soit $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ la suite des rentabilités obtenue en remplaçant les α_{2^m} ($m \in \mathbb{N}$) par -1 et α_0 par 4.

La série de cours $c_0, c_1, \dots, c_n, \dots$ qui correspond à la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est fortement battable à la vitesse 0.

Nous montrons maintenant :

- Il existe des suites de Martin-Löf qui peuvent garantir $c_n \in \mathbb{N}^*$.

Par construction, $c_0 = \sigma_0 = 4$, $\sigma_1 = \sigma_{2^0} = -1$, $\sigma_2 = \sigma_{2^1} = -1$, $\sigma_4 = \sigma_{2^2} = -1$, $\sigma_8 = \sigma_{2^3} = -1$, et $\sigma_3 = \sigma_5 = \sigma_6 = \sigma_7 = 1$. Le début de la série de cours est donc 4, 2, 1, 2, 1, 2, 4, 8, 4.

À partir du neuvième terme, c_n satisfait :

$$c_n = \begin{cases} \frac{c_{n-1}}{2} & n = 2^m \\ 2c_{n-1} & n \neq 2^m, \sigma_n = 1 \\ c_{n-1} & n \neq 2^m, \sigma_n = 0 \end{cases} \quad (4.56)$$

Puisque les c_{2^m} sont de plus en plus rares dans le temps, on peut garantir $c_n \in \mathbb{N}^*$ en modifiant un nombre fini de termes de la suite $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$.

- $c_0, c_1, \dots, c_n, \dots$ est fortement battable.

La stratégie $d(\cdot)$ qui consiste à tout vendre à chaque instant $2^m - 1$ et tout acheter à chaque instant $2^m + 1$ refuse uniquement les baisses de cours correspondant aux σ_{2^m} . La différence de gains entre $d(\cdot)$ et la "buy and hold" est :

$$r_n / r_n^* = 2^{|\log(n)|} \quad (4.57)$$

Où la fonction $|x|$ donne le plus grand entier qui ne dépasse pas x .

La limite de l'équation (4.57) étant $+\infty$, la stratégie $d(\cdot)$ bat fortement la série de cours $c_0, c_1, \dots, c_n, \dots$.

- $d(\cdot)$ bat $c_0, c_1, \dots, c_n, \dots$ à la vitesse 0.

Par définition, la vitesse de convergence est mesurée par l'équation suivante

$$\lim_{n \rightarrow \infty} \frac{\log(r_n / r_n^*)}{n^v} \quad (4.58)$$

En introduisant l'équation (4.57), on obtient,

$$\lim_{n \rightarrow \infty} \frac{\log(r_n/r_n^*)}{n^v} = \lim_{n \rightarrow \infty} \frac{\log(n)}{n^v} \quad (4.59)$$

Quand $n \rightarrow \infty$, quel que soit le paramètre $v > 0$, la fonction de puissance, n^v augmentera toujours plus vite que la fonction de logarithme, $\log(n)$. $\lim_{n \rightarrow \infty} \frac{\log(n)}{n^v}$ ne peut donc pas converger vers une valeur positive sauf si l'on impose un paramètre $v = 0$.

Ainsi, la stratégie $d(\cdot)$ bat fortement $c_0, c_1, \dots, c_n, \dots$ à la vitesse 0. ■

La proposition 4.6.1.1, élimine le schéma 4.3(a). Nous distinguons maintenant les schémas 4.3(b) et 4.3(c) en étudiant la possibilité de faiblement battre une série à la vitesse ∞ .

Proposition 4.6.1.2

Il existe des séries de cours faiblement battables à la vitesse $+\infty$.

Démonstration : Soit $\alpha_0, \alpha_1, \dots, \alpha_n, \dots$, une suite de Martin-Löf. On obtient la suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ en faisant les modifications suivantes :

1. Remplacer les termes pairs α_{2m} , $m \in \mathbb{N}$, par 2^m .
2. Modifier les termes α_{2m+1} comme suit :
Quand $m > 0$,

$$\sigma_{2m+1} = \begin{cases} \sum_{i=0}^{m-1} (-\sigma_{2i+1}) + 2^m & \text{si } \alpha_{2m+1} = 0 \\ \sum_{i=0}^{m-1} (-\sigma_{2i+1}) - 2^m & \text{si } \alpha_{2m+1} = 1 \end{cases}$$

Quand $m = 0$, $\sigma_1 = 1$ pour garantir que $c_1 = c_0\sigma_1 \in \mathbb{N}^*$.

La suite des rentabilités $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ ainsi construite est faiblement bat-
table à la vitesse $+\infty$.

- La stratégie $d(\cdot)$ utilise uniquement les σ_{2m} bat faiblement la “*buy and hold*”.

Selon la proposition 4.1.1.2, la différence de gains entre $d(\cdot)$ et la “*buy and hold*” est donnée par :

$$\lim_{n \rightarrow \infty} r_n/r_n^* = 2^{\lim_{n \rightarrow \infty} \sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} \sigma_{2i+1}} \quad (4.60)$$

Où la fonction $\lfloor \frac{n-1}{2} \rfloor$ désigne le plus grand nombre entier qui ne dépasse pas $\frac{n-1}{2}$.

Par construction, les termes impairs de la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ satis-
font :

$$\sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} \sigma_{2i+1} = \begin{cases} 2^{\lfloor \frac{n-1}{2} \rfloor} & \text{si } \alpha_{2\lfloor \frac{n-1}{2} \rfloor+1} = 0 \\ -2^{\lfloor \frac{n-1}{2} \rfloor} & \text{si } \alpha_{2\lfloor \frac{n-1}{2} \rfloor+1} = 1 \end{cases}$$

Selon l'équation (4.60),

$$\limsup_{n \rightarrow \infty} \log(r_n/r_n^*) = \limsup_{n \rightarrow \infty} 2^{|\frac{n-1}{2}|} = +\infty \quad (4.61)$$

et

$$\liminf_{n \rightarrow \infty} \log(r_n/r_n^*) = \liminf_{n \rightarrow \infty} -2^{|\frac{n-1}{2}|} = -\infty \quad (4.62)$$

$d(\cdot)$ bat donc faiblement la suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$.

– $d(\cdot)$ bat $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ à une vitesse infinie.

L'équation (4.63) calcule la vitesse de $d(\cdot)$ sur $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$.

$$\frac{\log(r_n/r_n^*)}{n^v} = \frac{(-1)^{\alpha_{2^{|\frac{n-1}{2}|+1}}} \times 2^{|\frac{n-1}{2}|}}{n^v} \quad (4.63)$$

Comme une fonction exponentielle augmente toujours plus vite qu'une fonction de puissance, quelle que soit la valeur de v , la \limsup de l'équation (4.63) est toujours ∞ .

La suite $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ est donc faiblement battue par $d(\cdot)$ à la vitesse $+\infty$. ■

La proposition 4.6.1.2 élimine le schéma 4.3(b). Nous avons donc trouvé la vraie relation entre la performance finale d'une stratégie donnée et sa vitesse de convergence : les deux critères jouent chacun un rôle dans l'évaluation des stratégies, aucun des deux ne peut être remplacé par l'autre.

4.6.2 Cas où la volatilité n'augmente pas dans le temps

Remarquons que la démonstration 4.6.1.2 est fondée sur une suite de rentabilités dont la volatilité tend vers ∞ . Or, sur les marchés financiers, cette situation est très peu réaliste. La volatilité des rentabilités financières est bornée, et ne peut pas augmenter infiniment.

Si la volatilité des cours n'augmente pas infiniment avec n , serait-il possible de trouver une borne supérieure à la vitesse de convergence ? La proposition suivante donne une réponse positive à cette question.

Définition "Suite à volatilité bornée" 4.6.2.1

Soit $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ une suite de rentabilités infinie, elle est dite "à volatilité bornée" si elle satisfait la condition suivante :

$$|\sigma_n| < \infty \quad (4.64)$$

où $|\sigma_n|$ désigne la valeur absolue de σ_n .

Proposition 4.6.2.2

Soit $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ une suite de rentabilités à volatilité bornée, $d(\cdot)$, une stratégie qui bat $\sigma_0, \sigma_1, \dots, \sigma_n, \dots$ à la vitesse v . Alors, pour toute $d(\cdot)$ calculable, nous avons $v \leq 1$.

Démonstration : Cette proposition est valable pour une raison simple : à l'instant n , une stratégie a eu au maximum n occasions de faire augmenter la différence de gains.

Comme la différence de gains réalisable à chaque instant dépend de $|\sigma_n|$, cette différence par instant ne peut pas croître avec n sur une suite à volatilité bornée.

La vitesse de convergence de $d(\cdot)$ est donc bornée à 1 sur la suite à volatilité bornée. ■

Par exemple, la série de cours $1, 2, 1, 2, 1, 2, \dots$ correspond à une suite de rentabilités à volatilité bornée. De toute évidence, cette série est fortement battable par la stratégie "acheter aux termes pairs et vendre aux termes impairs" à la vitesse 1. La proposition 4.6.2.2 implique que cette vitesse est le maximum qu'on puisse atteindre sur la série $1, 2, 1, 2, 1, 2, \dots$.

Avant l'introduction de la vitesse de convergence, les stratégies d'investissement sont uniquement évaluées selon leur performance à long terme. Une stratégie qui peut fortement battre une série de cours donnée est intuitivement considérée comme "meilleure" qu'une stratégie qui ne bat que faiblement ce même cours.

Or, la valeur du temps est centrale en théorie financière. Fortement battre une série à une vitesse insignifiante n'est pas toujours un meilleur choix que battre faiblement à une grande vitesse.

L'objectif de cette section était de mettre en avant la vitesse de convergence pour évaluer les stratégies d'investissement. Après en avoir défini les termes, nous avons étudié la relation entre la performance finale d'une stratégie donnée et sa vitesse de convergence. Deux résultats émergent de nos analyses théoriques :

- La vitesse et la performance finale sont deux critères complémentaires. Aucun des deux ne peut être remplacé par l'autre.
- Face à une suite à volatilité bornée, la vitesse de convergence d'une stratégie donnée ne dépasse pas 1.

Conclusion

Ce chapitre propose une réflexion autour de la possibilité de “battre le marché” en finance. Inspirée par la définition de suites aléatoires au sens de Schnorr (1971), nous définissons les cours imbattables par ceux qui ne sont battus par aucune stratégie calculable. Ce faisant, nous avons proposé un nouveau cadre théorique (que la théorie des probabilités) pour modéliser et évaluer les stratégies d’investissement.

Face à une série de cours donnée, nous proposons deux critères pour évaluer une stratégie de type “*market timing*” :

1. la performance de $d(\cdot)$ à long terme, mesurée par la différence de gains entre $d(\cdot)$ et la “*buy and hold*”.
2. la vitesse à laquelle $d(\cdot)$ atteint sa performance finale, mesurée par la vitesse de convergence de cette différence de gains.

La relation démontrée entre ces deux critères confirme les intuitions des professionnels : les meilleures stratégies en termes de performance ne sont pas toujours les plus rapides. Certaines stratégies sont tellement lentes à converger qu’on préfère les arrêter malgré leur performance à long terme.

Du point de vue des conditions d’échange, les stratégies calculables sont divisées en deux catégories : (1) les stratégies du type “tout acheter et tout vendre”, nommées “stratégies simples” (2) et les stratégies qui peuvent varier la mise à chaque intervention, dénommées “stratégies généralisées”. Il est montré que la modification de condition d’échange n’est pas toujours neutre sur la performance des stratégies. Plus précisément, les stratégies généralisées peuvent battre plus de suites que les simples.

Ce phénomène, issu de la dominance des stratégies géométriques, suggère l’usage des stratégies “tout acheter et tout vendre” dans les tests d’efficience. En effet, si les stratégies géométriques peuvent battre la “*buy and hold*” sur des séries de rentabilités imprévisibles, alors “battre le marché” avec une stratégie généralisée ne rejeterait plus directement l’efficience des marchés financiers.

Nous distinguons également des séries battables en présence d’une politique de blocage de celles qui le sont sans seuil de blocage. Cette différence aide à étudier l’impact des “*stop loss rules*”, une politique d’investissement qui impose la vente du titre quand la perte accumulée atteint un certain seuil et qui impose l’achat quand le cours aura suffisamment augmenté, sur la performance des stratégies à long terme.

Comme nous l’avons montré dans la proposition 4.3.0.3 (page 208), sous cer-

taines conditions de convergence, quelle que soit la valeur du seuil considéré, une stratégie pouvant battre une série de rentabilités donnée passera toujours à un moment critique à partir duquel celle-ci ne sera plus jamais bloquée.

Ainsi, même en présence des “*stop loss rules*”, une stratégie efficace ¹³ ne s’arrête qu’un nombre fini de fois, car elle passera à un moment critique à partir duquel elle ne sera plus jamais bloquée. La performance finale de cette stratégie n’est donc pas influencée par ces arrêts temporaires causés par le hasard.

Deux conséquences découlent de cette analyse :

- les “*stop loss rules*” ne sont pas responsables des mauvaises performances à long terme.
- la valeur du seuil de blocage ne semble pas avoir d’influence directe sur la performance finale des stratégies.

Bien sûr, ces conclusions ne sont valables qu’avec un horizon d’investissement infini, ce qui est irréaliste en pratique. Cependant, comme beaucoup de travaux théoriques en finance, on suppose que ces propriétés peuvent être observées quand la durée d’investissement est suffisamment longue.

La vitesse de convergence est aussi un critère important dans l’évaluation des stratégies d’investissement. Bien souvent, on ne cherche pas à réaliser la meilleure performance à long terme, mais une performance satisfaisante et rapide. Dans ce chapitre, la vitesse des stratégies est prise en compte par la moyenne géométrique de la différence de gains. Si une stratégie bat un cours à la vitesse “ $v = 1$ ”, cela signifie que celle-ci peut engendrer, en moyenne, une différence de gains positive à chaque instant passé. “ $v > 1$ ” implique que non seulement la stratégie engendre, en moyenne, une différence de gains positive à chaque instant passé, mais également que la différence réalisée par instant croît dans le temps. “ $v < 1$ ” indique que les occasions pour que la stratégie surperforme le marché deviennent de plus en plus rares avec le temps.

Une limite de cette étude se trouve dans l’absence de coûts de transaction. Cette absence est due au fait que la prise en compte de ces frais augmenterait radicalement la complexité des fonctions de richesse r_n . Nous pouvons peut être contourner cette difficulté en restant modéré dans nos conclusions : une suite battable en théorie ne l’est pas nécessairement en pratique. En revanche, une suite **imbattable** en théorie doit l’être également en pratique.

13. Au sens de pouvoir “battre le marché”.

CONCLUSION GÉNÉRALE

Cette thèse s'interroge sur les notions d'aléa et de régularité des variations boursières. Elle cherche à comprendre dans quelle mesure les marchés financiers sont imbattables ? Que révèlent les observations empiriques ? Comment faut-il les interpréter ? Telles sont les questions que nous avons essayé d'éclaircir au travers de quatre essais successifs ici présentés.

Les rentabilités excessives sont imprévisibles au sens où leur espérance conditionnelle est toujours égale à 0. Si cette proposition théorique est compatible avec les trois approches théoriques (i.e. efficacité informationnelle, finance comportementale, et approche conventionnaliste), sa négation semble être, pour le moment, la seule façon de falsifier l'efficacité informationnelle. En présence du problème de l'hypothèse jointe, tester l'absence de rentabilités excessives revient à vérifier l'impossibilité de "*market timing*". Une telle impossibilité est donc la propriété commune des séries de rentabilités mise en avant dans cette thèse.

Pour appuyer cette conclusion, des tests statistiques sont conduits pour caractériser les rentabilités successives. Bien que certaines structures, parfois rassemblées sous l'intitulé "faits stylisés", sont constatées dans leurs moments d'ordre supérieur, celles-ci ne permettent pas de battre la stratégie "*buy and hold*". Car les modèles économétriques fondés sur ces structures s'avèrent peu performants dans leur prédiction du sens des variations futures.

Sensibles uniquement aux régularités pour lesquelles ils sont définis, les tests statistiques pourraient rester aveugles à des régularités financières cachées derrière l'apparente complexité des dynamiques boursières. L'introduction d'une nouvelle approche empirique pourrait révéler des structures inconnues jusqu'à présent.

À cet effet, l'approche algorithmique, qui utilise les outils de compression pour détecter des régularités financières, est introduite pour estimer le degré d'aléa des rentabilités successives. Comme le prédit la théorie de l'efficacité, les résultats constatés à l'issue d'une telle opération restent en accord avec l'absence des rentabilités excessives.

Validée aussi bien par les théories financières que par les faits empiriques, l'impossibilité de "*market timing*" est, selon nous, une propriété commune des séries des rentabilités boursières. Nous définissons donc les cours imbattables dans ce sens spécifique en modélisant les principes d'investissement par les fonctions calculables. Cette définition propose un nouveau cadre théorique pour comparer et étudier des stratégies financières. Ce nouveau cadre issu de la théorie de la calculabilité nous permet d'étudier une série de rentabilités particulière mais non une famille de suites probables comme ce qui est proposé par les approches statistiques.

Les principaux résultats obtenus dans cette thèse peuvent être résumés comme suit :

- Il existe des régularités et une certaine forme de prévisibilité dans les séries de rentabilités successives. En revanche, ces régularités constatées ne permettent pas de battre le marché.

Du point de vue théorique, l'absence des *rentabilités excessives* n'implique en aucun cas l'indépendance des variations de prix successives. L'existence de certaines structures dans les séries de rentabilités demeure donc tout à fait compatible avec les théories financières étudiées. Cependant, ces structures doivent rester inexploitable au sens du "*market timing*", car quel que soit le modèle d'équilibre retenu, l'espérance des rentabilités d'équilibre est toujours positive. La meilleure stratégie qu'on peut concevoir est donc de garder les titres à long terme, comme le propose la stratégie "*buy and hold*". Sur le plan empirique, bien que les rentabilités successives ne soient pas *i.i.d.*, il est pour autant très difficile (voire impossible) de battre la "*buy and hold*" à long terme. Les modèles économétriques sont peu efficaces à prédire le sens des variations futures. En ce sens le marché est "efficient" dans la mesure où ses fluctuations semblent bien découler de la survenue de nouvelles imprévues, ceci n'ayant pas avoir directement avec la pleine révélation dans les prix (les trois corps théoriques sont acceptables à cet égard).

- On introduit le concept de complexité algorithmique en finance qui est mis en oeuvre pour mesurer le degré de prévisibilité et d'aléa des séries de rentabilités.

La complexité de Kolmogorov d'une série s correspond, pour certaines définitions, à la taille du plus petit programme capable d'engendrer s . Sur la base d'une telle définition, on discute du degré d'aléa dans les séries financières. Autrement dit, on introduit la notion de complexité algorithmique en finance, ainsi que des procédures opérationnelles pour en estimer une valeur.

Les séries de rentabilités journalières sont complexes au sens où elles combinent des structures régulières qui peuvent, par effacement progressif ré-

véler un “noyau” radicalement aléatoire. On montre qu’après un processus d’effacement itératif de ces régularités, il demeure un “bruit” qu’on pourrait croire uniforme, mais qui peut être compressé dans une certaine mesure, ce qui indique l’existence de régularités plus profondes.

Néanmoins, étant donné la valeur relativement faible du taux de compression obtenu après l’effacement des régularités apparentes (0,49%), la structure restant, en admettant sa présence, semble peu exploitable sur le plan pratique.

Malgré leur valeur plus élevée, les taux de compression affichés pour les séries de rentabilités “*tick-by-tick*” nous semblent principalement liés à des éléments micro-structuraux et reste compatible avec l’impossibilité de battre la stratégie “*buy and hold*”.

- Ce travail enfin propose une définition mathématique des “cours imbattables”. L’impossibilité de “*market timing*”, implication et probablement la seule implication de l’efficience informationnelle au regard des dynamiques boursières est mise en évidence pour les chercheurs de toutes les disciplines. Définir les cours imbattables en tant qu’objet mathématique permet également d’étudier l’ensemble des cours boursiers, mais non des réalisations spécifiques du modèle proposé, tels que les indices Dow Jones, CRSP, ou CAC40.

En étudiant l’impact de certaines politiques d’investissement sur la possibilité de battre un cours donné, on obtient 3 conclusions générales.

- En imposant une politique de blocage aux stratégies calculables, on propose un argument en faveur des “*stop loss rules*”. Si cette pratique n’améliore pas la performance des stratégies considérées, elle n’est pas non plus responsable de leur échec à long terme.
- En intégrant les stratégies généralisées, i.e., celles qui peuvent varier la mise à chaque intervention, on démontre que certains cours imbattables par des stratégies “tout acheter et tout vendre” sont battues par ces stratégies généralisées.
- La vitesse à laquelle une stratégie engendre les différences de gains n’a pas d’incidence sur sa performance finale. La vitesse et la performance sont deux critères relativement indépendants pour évaluer une stratégie donnée.

En dépit de son enjeu théorique et ses perspectives séduisantes, l’approche algorithmique présente tout de même des limites méthodologiques. Nos détecteurs de régularités, i.e., les algorithmes de compression, sont pour l’instant “empruntés” à la compression de texte en informatique. Ceux-ci ne sont donc pas conçus pour les études des séries financières. En concevant des algorithmes spécialement destinés aux régularités financières, de nouvelles structures pourraient être révélées dans les dynamiques boursières.

CONCLUSION GENERALE

Enfin, comme nous l'avons défendu dans le deuxième chapitre, certaines régularités, notamment celles cachées dans les moments d'ordre supérieur, n'agissent pas sur le sens des variations futures. Bien que ces structures permettent de compresser les séries étudiées, elles n'ont aucun impact sur leur niveau d'efficience. Les taux de compression obtenus nécessitent donc une analyse ultérieure sur la nature des régularités révélées, comme nous l'avons fait pour les données *"tick-by-tick"* observées sur les marchés internationaux.

En ce sens, le degré d'aléa proposé dans cette étude n'est pas un indicateur direct de l'efficience informationnelle. Les recherches futures pourraient donc améliorer les algorithmes de compression en orientant leur fonctionnement vers la recherche des seules régularités qui sont reliées au sens des variations futures.

ANNEXE

A. EFFICIENCE DES MARCHÉS FINANCIERS DANS UN CADRE D'INFORMATION IMPARFAITE ET D'AGENTS HÉTÉROGÈNES

L'information parfaite et l'homogénéité des agents sont toutes deux les conditions requises d'un marché parfait, **bien dissociées** de l'efficience informationnelle (Ball 1989). Or, il semblerait que peu d'auteurs à l'époque aient remarqué la précision de Fama citée précédemment, et le débat sur l'efficience informationnelle a donc commencé par le rejet de ces deux conditions jugées irréalistes.

Par exemple, sous l'hypothèse d'homogénéité, Grossman et Stiglitz (1980) s'interrogent sur la motivation des investisseurs à spéculer sur les marchés financiers. Selon ces auteurs, si toutes les informations disponibles sont instantanément intégrées dans les prix de transactions, alors, tous les agents homogènes devraient avoir les mêmes évaluations des actifs, ils n'ont donc aucune raison à échanger leurs titres. Ceci entre en contradiction flagrante avec les volumes de transactions qu'on observe tous les jours sur les places boursières.

Black (1986) explique les échanges boursiers par la présence de "bruits" ("*noise*") sur les marchés financiers. En effet, sur un marché bruité, les investisseurs n'ont pas d'accès aux mêmes informations, ce qui engendre une divergence d'opinion concernant l'évaluation des titres. Les investisseurs qui sous-évaluent un titre vont le vendre à ceux qui le sur-évaluent, on observe ainsi les transactions dynamiques sur les marchés.

Or, selon Aumann (1976), Tirole (1982) et Milgrom et Stokey (1982), la présence de "bruits" ne suffit pas à elle seule à engendrer les transactions financières. Si les investisseurs sont homogènes dans le sens où ils interprètent les informations fondamentales avec le même modèle d'évaluation, alors, en consultant les ordres émis par les autres investisseurs, les individus *non informés* peuvent extraire de l'information des signaux reçus par les agents *informés*. Ce processus d'extraction, fondé sur la condition d'homogénéité des agents, finit par éliminer les "bruits" sur le marché, et ce faisant, réduit à néant la motivation à échanger.

Ces analyses affinées ont apporté un nouvel éclairage sur la notion d'efficience :

pour que les investisseurs soient incités à échanger leurs titres, leur rationalité ne peut qu'être définie dans un cadre d'hétérogénéité, notamment en ce qui concerne le choix du modèle d'évaluation.

La théorie de l'efficience informationnelle s'est ainsi adaptée en intégrant dans les postulats de base l'hétérogénéité des agents économiques. L'efficience devient alors un phénomène collectif émergé des interactions entre l'ensemble des investisseurs qui ne partage pas la même conviction sur le "bon modèle" d'évaluation.

Par exemple, selon Rubinstein (1975), un marché est "efficient vis à vis de l'information A ", si l'ajustement de prix provoqué par A est identique à celui d'un marché où son arrivée est à la connaissance de chaque investisseur. Beaver (1981) définit l'efficience informationnelle par le fait que toutes les informations soient interprétées comme si leur arrivée était déjà à la connaissance de chaque investisseur. Et Latham (1986) qualifie d'efficient, un marché où l'annonce d'une information ne modifierait ni les prix d'équilibre ni les portefeuilles d'équilibre.

Les travaux théoriques cités précédemment, visent avant tout à préciser la façon dont les informations fondamentales sont intégrées sur un marché efficient. Au lieu de les interpréter comme une donnée homogène traitée identiquement par chaque investisseur rationnel, les "conventions de marché" au sens de Fama (1970) ne peuvent qu'être le résultat des interactions entre les agents hétérogènes. Autrement dit, sur un marché efficient, sans que tous les investisseurs partagent la même conviction sur la valeur d'un titre, l'ensemble de leurs interactions fait émerger le prix d'équilibre qui intègre toutes les informations disponibles. L'existence de ce prix d'équilibre, nommé "*consensus belief*", a été rigoureusement démontrée par Verrecchia (1975).

En revanche, d'un point de vue empirique, ces contributions théoriques n'engendrent pas d'implications véritablement testables à la théorie de l'efficience. C'est probablement pour cela que la version de Fama (1970) reste la définition d'efficience la plus souvent citée par les travaux empiriques (Ball 1989).

B. “CROSS SECTION TESTS” SUR LA VALIDITÉ DU CAPM

A l’opposé de l’analyse “*time series*”, l’analyse “*cross section*” néglige la notion de temps dans les données observées. Par exemple, pour connaître le taux d’obésité prévalant dans un pays, on peut tirer au hasard un échantillon de 1000 individus, mesurer leurs poids et taille, et calculer le taux d’obésité. Cette étude est de nature “*cross section*” dans la mesure où l’ordre d’arrivée des individus est complètement négligé.

Un test “cross section” qui évalue la validité du CAPM se déroule en 6 étapes :

1. Choisir un ensemble de titres à analyser. Par exemple, dans le travail de Blume (1970), l’auteur a choisi 251 titres financiers cotés sur le New York Stock Exchange. Puis, fixer une période d’observations qui doit être commune pour chaque actif à analyser. (en ce cas, du 30/12/1926 au 30/12/1960).
2. Estimer le $\beta_{i,m}$ empirique de chaque actif i , en effectuant une régression linéaire entre les rentabilités de chacun d’eux et celles du marché.
3. Classer l’ensemble des actifs dans 10 portefeuilles selon leur niveau de β . Le premier portefeuille est composé des 10% des actifs dont les $\beta_{i,m}$ sont les moins élevés, et le dernier portefeuille est composé par les 10% des actifs dont les $\beta_{i,m}$ sont les plus élevés.
4. Pour chaque portefeuille composé à l’étape précédente, calculer son β et la moyenne de ses rentabilités réalisées. Les moyennes des rentabilités seront considérées comme les rentabilités espérées des portefeuilles correspondants.
5. On obtient 10 rentabilités espérées avec leur β correspondant.
6. Effectuer une régression linéaire entre les rentabilités espérées et les β des 10 portefeuilles. Selon l’équation 1.2, la pente obtenue dans cette régression (notée θ) est égale à la prime du risque $E_m - R_f$.

Fama et MacBeth (1973) proposent de séparer la période d’observations en plusieurs sous-périodes de longueur identique (noté n le nombre de sous-périodes obtenues). En appliquant la technique de Blume (1970) à chaque sous-période, on tient n observations de la pente θ . Un test statistique peut ainsi être conduit pour vérifier si θ est significativement différent de sa valeur théorique.

La conclusion commune des “cross section tests” se trouve dans l’infériorité

significative de θ par rapport à la prime du marché théorique $E_m - R_f$ ¹, où la rentabilité du marché E_m est estimée par celle de l'ensemble des actions américaines et R_f par celle des emprunts obligataires émis par l'état américain.

1. Cette infériorité a également été confirmée par Black, Jensen, et Scholes (1972) et Fama et French (1992)

C. DESCRIPTION DU TEST GRS

Les observations de Black, Jensen, et Scholes (1972) étant de nature “ex post”, Gibbons, Ross, et Shanken (1989) (désormais GRS) proposent un test statistique pour vérifier le rejet du CAPM. Ce test de GRS se procède en 5 étapes :

1. Choisir le portefeuille du marché. Le travail de GRS utilise l'indice CRSP¹ observé pendant la période 01/1931 – 12/1965.
2. Soit p , le portefeuille à analyser, R_p les rentabilités observées de p , $r_p = R_p - R_f$ les rentabilités excessives de p , $\sigma(r_p)$ l'écart-type des rentabilités excessives² et $\theta = \frac{E(r_p)}{\sigma(r_p)}$ le ration risque-rentabilité offert par p .
3. Choisir librement N actifs se trouvant dans le portefeuille p . La validité de CAPM dépend donc de la relation entre les rentabilités de ces N actifs et celle du portefeuille p .
Pour chaque actif choisi, calculer le même paramètre $\theta_i = \frac{E(r_i)}{\sigma(r_i)}$. Noté θ^* , le θ_i le plus élevé parmi les N observations.
4. Avec un F test, nous pouvons savoir si

$$W = \left[\frac{\sqrt{1 + \theta^{*2}}}{\sqrt{1 + \theta_p^2}} \right]^2 - 1 \quad (\text{C.1})$$

est significativement différent de 0. En cas de réponse positive, le CAPM est rejeté.

Le test GRS s'interprète graphiquement de la manière suivante :

- Dans la figure C.1, le paramètre θ_p du portefeuille p (resp. actif i) désigne la rentabilité offerte par p (resp. actif i) pour chaque unité de risque subi. θ^* représente la meilleure rentabilité qu'on puisse obtenir pour une unité de risque subi.
- Le F test sur W examine donc si la distance od est significativement différente de oc .

1. pondéré avec le même poids pour chaque actif.

2. Le test GRS travaille avec les données mensuelles, ils ont donc 409 observations de r_p .

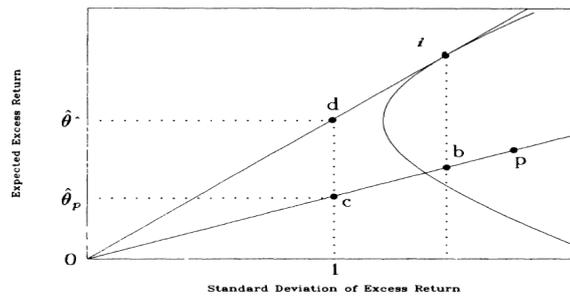


Fig. C.1: L'illustration graphique du test GRS.

Ici, l'axe X représente le niveau du risque mesuré en terme de l'écart-type, l'axe Y la rentabilité espérée. Source : la figure 1a dans l'article Gibbons, Ross, et Shanken (1989).

D. PRÉSENTATION DES TROIS MODÈLES COMPORTEMENTAUX

Nous présentons ici 3 modèles comportementaux pour illustrer l'usage des biais psychologiques dans l'explication des rentabilités excessives.

1. Le modèle de Barberis, Shleifer, et Vishny (1998) (désormais BSV) est fondé sur le biais de "*conservatism*". Selon BSV, les informations récentes, par exemple une baisse de dividende, risquent d'être sous interprétées par le marché, car les agents peuvent croire à tort que cette baisse est temporaire et refusent de réviser leur anticipation initiale. C'est uniquement quand cette croyance sera démentie par les informations futures que les investisseurs réajusteront leur anticipation de prix vers le bas. L'adaptation tardive du modèle de prévision engendra une auto-corrélation positive des rentabilités excessives à court terme.

De plus, en observant **par hasard** les informations successives qui font toutes augmenter (ou baisser) la valeur de l'actif, les investisseurs les interprètent souvent comme une tendance irréversible (i.e "*representativeness heuristic*") et sur-réagissent par rapport à l'importance des informations observées. Quand cette tendance artificielle se verra corrigée par les informations futures, le prix de l'actif retournera vers sa valeur fondamentale. On observe ainsi une auto-corrélation négative dans les rentabilités à long terme, qui peut également aider à la prévision des rentabilités excessives.

2. Daniel, Hirshleifer, et Subrahmanyam (1998) (désormais DHS) prédisent les mêmes corrélations en s'appuyant sur le biais "*over-confidence*". Selon DHS, les agents réagissent en fonction de leurs informations privées. Un investisseur qui reçoit un signal positif, va acheter le titre concerné pour profiter de cette information privée. Après l'achat, l'agent a une tendance à sur-pondérer les informations publiques validant son investissement, et à sous-pondérer celles qui le contredisent. Autrement dit, en moyenne, les informations publiques après l'achat seront traitées dans le sens du signal privé (Daniel, Hirshleifer, et Subrahmanyam 1998, p.1842). Ceci entraîne une sur-réaction du marché par rapport à l'information privée. On observe ainsi une auto-corrélation positive des rentabilités à court terme. En revanche, sur le long terme, l'accumulation des informations publiques ramènera le prix de l'actif vers sa valeur fondamentale, et ce retour de valeur engendrera une

auto-corrélation négative des rentabilités à long terme.

3. Selon Hong et Stein (1999) (désormais HS), le marché est composé de deux types d'investisseurs : les "*newswatchers*" qui répondent aux informations privées et les "*momentum traders*" qui achètent (resp. vendent) quand les prix augmentent (resp. baissent). Les auteurs supposent ensuite que les informations fondamentales se propagent de manière progressive parmi les "*newswatchers*". Elles ne sont donc incorporées que progressivement dans les prix de transaction (sous-réaction), et ceci entraîne une auto-corrélation positive dans les rentabilités à court terme. Cette corrélation déclenchera les achats ou les ventes des "*momentum traders*", qui feront écarter les prix des actifs de leur valeurs fondamentales (sur-réaction). Comme dans les modèles précédents, la sur-réaction du marché implique le retour de valeur qui, à son tour, provoquera une corrélation négative entre les rentabilités à long terme.

En comparant ces trois modèles, on constate que la finance comportementale donne de différentes explications aux mêmes faits empiriques : par exemple, en comparant le modèle BSV et le modèle DHS, on constate que l'auto-corrélation positive des rentabilités à court terme peut aussi bien être provoquée par la sous-réaction que par la sur-réaction des investisseurs.

E. ORIGINE THÉORIQUE DU TEST DE VOLATILITÉ ET DU PHÉNOMÈNE DE RETOUR À LA MOYENNE

Les tests de volatilité organisés sous deux hypothèses théoriques :

1. le marché est efficient
2. les valeurs fondamentales sont évaluées selon le modèle de dividendes.

Sous ces hypothèses, le prix d'un actif donné est la meilleure estimation du prix futur (dividende ajusté), c'est-à-dire que,

$$p_t = \alpha \times E((D_{t+1} + p_{t+1})|\Omega_t) \quad (\text{E.1})$$

Où p_t et p_{t+1} désignent respectivement le prix à l'instant t et $t + 1$. D_{t+1} représente le dividende par action distribuée à l'instant $t + 1$, Ω_t signifie toutes les informations disponibles à l'instant t , et α le taux de rentabilité sans risque, aussi appelé le facteur d'actualisation .

En introduisant

$$e_{t+1} = p_{t+1} + D_{t+1} - E(D_{t+1} + p_{t+1}|\Omega_t), \quad (\text{E.2})$$

l'équation (E.1) se transforme en :

$$p_t = \alpha(D_{t+1} + p_{t+1}) - \alpha \times e_{t+1} \quad (\text{E.3})$$

En remplaçant t par $t + i$ et en multipliant les deux cotés par α^i , la relation (E.3) devient,

$$\alpha^{i+1} \times \alpha^{i+1} D_{t+i+1} = \alpha^i p_{t+i} - \alpha^{i+1} p_{t+i+1} + \alpha^{i+1} \times e_{t+i+1} \quad (\text{E.4})$$

En prenant le $\sum_{i=0}^{\infty}$ de l'équation (E.4), on obtient

$$\sum_{i=0}^{\infty} \alpha^{i+1} \times D_{t+i+1} = p_t - \alpha^{\infty} p_{\infty} + \sum_{i=0}^{\infty} \alpha^{i+1} \times e_{t+i+1} \quad (\text{E.5})$$

En supposant $\alpha^{\infty} p_{\infty} = 0$, l'équation (E.5) devient,

$$\sum_{i=0}^{\infty} \alpha^{i+1} \times D_{t+i+1} = p_t + \sum_{i=0}^{\infty} \alpha^{i+1} \times e_{t+i+1} \quad (\text{E.6})$$

A remarquer que la partie gauche de l'équation (E.6) est la valeur fondamentale du titre à l'instant t selon le modèle de dividende. Elle est notée p^* et nommée "le prix rationnel **ex post**" par Shiller (1981).

Si p_t est indépendant de l'erreur d'anticipation, notée

$$x = p^* - p_t = \sum_{i=0}^{\infty} \alpha^{i+1} \times e_{t+i+1}$$

alors de l'équation (E.6) on obtient la relation

$$V(p^*) = V(p_t) + V(x) \quad (\text{E.7})$$

Où $V(\cdot)$ désigne la fonction de la variance. Comme les variances sont toujours positives, on a l'inégalité suivante :

$$V(p^*) \geq V(p_t) \quad (\text{E.8})$$

De plus, la violation de l'inégalité (E.8) implique le phénomène du retour de moyenne. En effet, tester la relation (E.8) revient à tester l'orthogonalité (l'indépendance) entre p_t et x (LeRoy et Porter 1981). $V(p_t) + V(x) > V(p^*)$ implique une corrélation négative entre p_t et x . Par définition, p_t est le prix initial p_0 plus la somme des rentabilités observées avant l'instant t , et x est la somme des rentabilités actualisées après cet instant, la corrélation négative entre les deux implique une corrélation négative entre la moyenne des rentabilités historiques et la moyenne des rentabilités futures.

A remarquer que le test de volatilité n'est pas identique aux tests d'auto-corrélation classique, ces derniers vérifient l'indépendance des rentabilités à court terme (journalières ou hebdomadaires), alors que le test de volatilité "teste l'orthogonalité entre une moyenne lisse des rentabilités historiques pendant plusieurs années et la moyenne similaire des rentabilités futures (LeRoy 1989, p.1601)¹".

1. Texte en anglais : "test the orthogonality of a smooth average of past returns over a period of years and a similar smooth average of future returns (LeRoy 1989, p.1601)."

F. STATISTIQUE BDS

Nous voulons étudier la chronique $\{X_1, X_2, \dots, X_n\}$ de taille n afin de déterminer si ses valeurs sont *indépendantes et identiquement distribuées*. A cet effet, on commence par se doter d'une distance ε .

Si les données étudiées sont bien *iid*, on note la probabilité $c_1(\varepsilon)$ pour que deux valeurs choisies au hasard aient une distance inférieure ou égale à ε :

$$c_1(\varepsilon) = Pr(|X_t - X_s| \leq \varepsilon) \quad (\text{F.1})$$

Plutôt que de raisonner point à point, on constitue des ensembles de paires de points issus de la chronique. Par exemple $\{(X_1, X_2), (X_2, X_3), \dots, (X_{n-1}, X_n)\}$ ou bien encore $\{(X_4, X_6), (X_5, X_7), \dots, (X_{n-2}, X_n)\}$.

Ce sous ensemble est constitué de $m = n - s - 1$ points consécutifs issus de la chronique. Dans le test BDS, ces sous-ensembles de paires de points sont des enveloppements de dimension m .

La probabilité pour que chaque paire de points dans l'enveloppement respecte la condition de distance se note $c_m(\varepsilon)$.

Si X_i est bien *iid*, alors :

$$c_m(\varepsilon) = c_1^m(\varepsilon) \quad (\text{F.2})$$

Autrement dit, sous H_0 , la probabilité pour que toutes les paires de points dans l'enveloppement respectent la condition de distance est égale au produit des probabilités qu'individuellement chacune de ces paires le fasse elle même.

La statistique BDS, notée $W(n, m, \varepsilon)$, sous H_0 , est asymptotiquement distribuée selon une $N(0, 1)$.

$$W(n, m, \varepsilon) = \sqrt{n} \frac{C(n, m, \varepsilon) - C(n, 1, \varepsilon)^m}{\hat{\sigma}(n, m, \varepsilon)} \quad (\text{F.3})$$

Le test BDS permet donc de tester H_0 contre une grande variété de dépendances incluant celles qui seraient linéaires et non linéaires. Ce n'est toutefois *pas un test direct* de présence de chaos ou de non linéarité, car on ne connaît pas la distribution de cette statistique sous ces hypothèses.

G. ORIGINALITÉ DES SUITES COMPRESSIBLES

Il y a au plus 2 programmes de longueur 1 et plus généralement au plus 2^i programmes de longueur i . Certaines suites de 0 et de 1 sont des programmes pour le langage universel de référence L . D'autres non. Il y a donc moins de :

$$2 + \dots + 2^{h-1} = 2^h - 2$$

programmes de longueur strictement inférieure à h .

Il y a par ailleurs 2^n suites de longueur n . La proportion de celles qui sont compressibles de plus de k digits est inférieure à

$$2^{n-k} - 2/2^n < 1/2^k$$

En prenant $k = 10$ on obtient qu'au plus 1/1024 est compressible de plus de 10 digits et qu'au plus 1/1048576 est compressible de plus de 20 digits.

H. COMPRESSION DE LA SUITE DE CHAMPERNOWNE

Voici le programme (écrit dans le langage “R”¹) qui engendre les digits de la suite de Champernowne correspondant au codage des entiers de 1 à 10000 en binaire mis les uns derrière les autres.

```
c < -0; j < -0; d < -0
for (n in 1 : 10000) {a < -n; k < -0;
while(a! = 0){k < -k + 1; d[k] < -a%%2; a < -(a - d[k])/2};
for (h in 1 : k) {j < -j + 1; c[j] < -d[k - h + 1]}
```

Cette suite a comme longueur 123631 digits, tandis que le programme qui a 132 caractères (soit 1056 digits). Nous obtenons donc un taux de compression de 99,145%.

Si nous allions plus loin dans la suite de Champernowne, le taux de compression serait meilleur : il tend vers 100% quand n tend vers l’infini.

1. <http://www.r-project.org/>

I. ANALYSES STATISTIQUES DE LA DISTRIBUTION $N(0, 1)$

Tab. I.1: Tests d'autocorrélation

Série	$\chi - deux$	deg. lib.	$p-value$
baseline	0.1838	1	0.668
	36.9157	36	0.4264

Interprétation : *On ne peut rejeter H_0 , les termes de la série sont indépendants*

Tab. I.2: Test BDS pour *baseline*, $m = \{2, 3\}$

ε	0.5012	1.002	1.5035	2.0047
$m = 2$	-0.2082	-0.2987	-0.5232	-0.7221
$p-value$	0.8351	0.7651	0.6009	0.4702
$m = 3$	0.8351	0.7651	0.6009	0.4702
$p-value$	0.9503	0.9803	0.9344	0.8600

Interprétation : *On ne peut rejeter H_0 , les termes de la série sont $i.i.d.$*

Tab. I.3: Tests de racine unitaire pour *baseline*

Test	Val. statistique	ordre	$p-value$
ADF	-32.7479***	31	0.01
PP	-178.52949***	16	0.01

Interprétation : *On peut rejeter H_0 : "La série est stationnaire"*

Cet ensemble de résultats était prévisible et n'est donné qu'à titre de point de repère pour la suite de la démonstration.

J. DIFFÉRENCE ENTRE UN JEU ÉQUITABLE ET UN PROCESSUS INDÉPENDANT

Proposition J.1

Il existe des processus jeu équitable qui ne sont pas indépendants.

Démonstration : Soit Y_t un processus discret respectant :

$$Y_{t+1} = Y_t + \sigma_{t+1}U_{t+1} \quad (\text{J.1})$$

Où U_t est *i.i.d* selon une loi normale centrée réduite. σ_t est indépendant de U_t , mais dépendant des valeurs réalisées de Y_t en période t et $t - 1$. Sans perdre de généralité supposons $\sigma_t = 1.5 * |Y_t - Y_{t-1}|$.

Le processus $\Delta Y_t = Y_{t+1} - Y_t$ est un jeu équitable puisque :

$$\begin{aligned} E(Y_{t+1}|Y_t, Y_{t-1} \dots Y_1) &= E(Y_t|Y_t, Y_{t-1} \dots Y_1) + E(\sigma_t U_t|Y_t, Y_{t-1} \dots Y_1) \quad (\text{J.2}) \\ &= Y_t + E(\sigma_t|Y_t, Y_{t-1} \dots Y_1) * E(U_t) \\ &= Y_t \end{aligned}$$

Or, ΔY_t n'est pas un processus indépendant car ses variances sont auto-corrélées :

$$\begin{aligned} E(\Delta Y_t^2|Y_t, Y_{t-1} \dots Y_1) &= E(Y_{t+1} - Y_t)^2|Y_t, Y_{t-1} \dots Y_1 \quad (\text{J.3}) \\ &= \sigma_t^2 V(U_t) = |Y_t - Y_{t-1}|^2 * 1.25 \\ &= (\Delta Y_{t-1})^2 * 1.25 \end{aligned}$$

■

K. QUELQUES NOTIONS ÉLÉMENTAIRES EN THÉORIE DE LA CALCULABILITÉ

Pour présenter rigoureusement les définitions de hasard en informatique, il convient d'étudier, ici, quelques notions élémentaires au regard de la complexité.

Commençons par quelques précisions de vocabulaire :

- En informatique, le terme “**programme**” désigne une séquence d'instructions qui peut être exécutée par un ordinateur. Il est souvent écrit dans un langage de programmation qui peut être compilé ou interprété par les ordinateurs.
- **Un algorithme** est une suite d'opérations logiques destinées à résoudre un problème donné. Il est souvent considéré comme le plan de programmation. Un algorithme exprimé dans un langage de programmation est ce qu'on appelle un programme.
- Le mot “**automate**” engendre souvent des confusions, car il signifie **un programme** exécutable par un ordinateur et non la machine qui l'exécute comme peut laisser l'entendre le langage commun.
- **Une procédure effective** est une solution à un problème donné correspondant à un programme exécutable par un ordinateur. Sans doute, cette définition nécessite davantage de précision sur le type de machine qu'on peut utiliser ce qui fera l'objet de discussion dans la suite de cette section.
- **Une machine de Turing** est une forme d'automate spécifique. Proposée par Turing (1936) celle-ci est devenue la référence des procédures effectives en théorie de la calculabilité. La machine de Turing est présentée en détail avec un exemple concret dans la suite de cette section, car c'est sur sa capacité de calcul que se base la notion de hasard en informatique.

D'abord, une machine de Turing est un automate composé de 4 éléments essentiels :

1. Une mémoire infinie sous forme de “ruban” divisé en cases de taille identique. Chaque case peut contenir un caractère provenant d'un alphabet fixé. Par exemple, pour une machine de Turing qui traite uniquement les textes en français, son alphabet comporte les 26 lettres de la langue française ¹, et chaque case du ruban ne peut contenir qu'une seule lettre.

1. plus éventuellement le symbole représentant le vide.

2. Une tête de lecture (“*read head*” ou “*tape head*”) qui peut se déplacer vers la gauche ou la droite du ruban selon les instructions données par le programme.
3. Un ensemble fini d'états parmi lesquels on distingue un état initial et un ensemble d'états acceptateurs.
4. Une fonction de transition qui indique pour chaque état de la machine et caractère lu par la tête de lecture :
 - l'état suivant de la machine de Turing,
 - le caractère à utiliser pour remplacer le caractère lu.
 - le déplacement de la tête de lecture (gauche ou droite).

Par exemple, la machine de Turing e qui calcule pour chaque nombre n , le montant $n + 1$, peut prendre la forme suivante :

- l'alphabet $\Gamma = \{0, 1\}$,
- la machine se trouve éventuellement dans 2 états possibles, l'état initial q_0 et l'état d'arrêt q_1 .
- le symbole vide est désigné par $\#$.
- à l'état initial, la tête de lecture se trouve au dernier symbole non vide du ruban.
- la fonction de transition est définie par le tableau suivant :

Ancien état	Symbole lu	Symbole écrit	Mouvement	Nouvel état
q_0	0	1	Gauche	q_1
	1	0	Gauche	q_0
	$\#$	0	Gauche	q_1

Pour illustrer le fonctionnement de la machine Turing e , supposons que son ruban d'entrée se présente comme $\#101\#$ (i.e représentation binaire du nombre 5), alors, l'état initial de la machine est décrit par la figure K.1 : la tête de lecture est positionnée au dernier symbole non vide du ruban, e se trouve en état q_0 . Selon

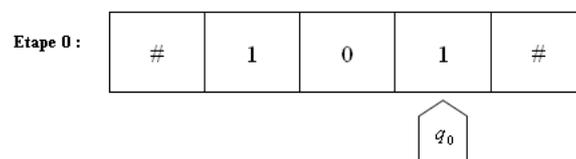


Fig. K.1: L'état initial de la machine de Turing e

la fonction de transition, en lisant le symbole “1”, la tête de lecture doit écrire le symbole 0 (à la place de “1”), puis faire un pas vers le gauche. L'état de la machine reste q_0 . Cette opération est présentée par la figure K.2. A cette nouvelle position, la tête de lecture lit le symbole “0”. Selon la fonction de transition, “1”

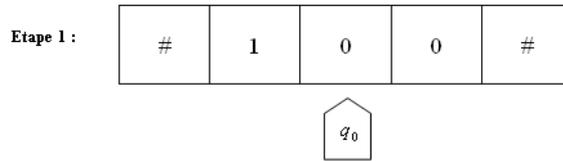


Fig. K.2: La première étape de calcul.

est écrit à la place de “0”, la tête de lecture se déplace d’un pas vers le gauche, et la machine de Turing e se trouve dans son état d’arrêt q_1 . Comme nous pouvons le constater dans la figure K.3, après ces calculs, le ruban devient $\#110\#$ qui est la représentation binaire du nombre 6.

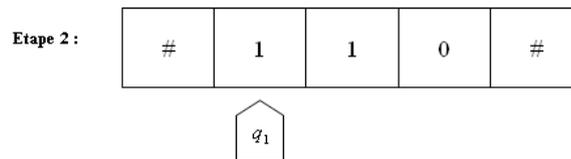


Fig. K.3: La deuxième étape de calcul.

K.1 Thèse de Turing-Church et machine de Turing universelle

La thèse de Turing-Church affirme que

Thèse de Turing-Church K.1.0.1

Tout problème soluble par une procédure effective l’est également par une machine de Turing.

Autrement dit, bien que les machines de Turing ne sont qu’une catégorie spécifique des automates, elles peuvent représenter l’ensemble d’automates en terme de calculabilité, car elles sont capables de résoudre tous les problèmes solubles par les procédures effectives.

Certes, cette proposition reste une thèse et non un théorème, il n’existe donc aucune démonstration formelle de sa validité. Or, comme tous les théorèmes prouvés après la thèse de Turing-Church sont parfaitement cohérents avec cette thèse

et aucun contre exemple n'est constaté jusqu'à nos jours, la thèse de Turing-Church est souvent admise en calculabilité. Par conséquent, représenter les procédures effectives par les machines de Turing est une pratique fréquemment rencontrée en informatique fondamentale.

Turing (1936) a démontré l'existence d'une machine de Turing ultime qui peut simuler le fonctionnement n'importe quelle machine de Turing. Plus précisément, il existe une machine de Turing M telle que si on lui fournit la description d'une machine de Turing quelconque M' et un mot d'entrée ω , M simulera l'exécution de M' sur ω . Une telle machine de Turing est appelée *machine de Turing universelle*.

En admettant la thèse de Turing-Church, la machine de Turing universelle est donc un automate qui peut résoudre tous les problèmes solubles par une procédure effective.

K.2 Ensemble récursif et ensemble récursivement énumérable

Dans cette section, on va étudier un problème classique en informatique : l'acceptation d'un ensemble dénombrable par une machine de Turing. Pour préciser, un ensemble infini Ω est dit *dénombrable* s'il existe une bijection entre Ω et l'ensemble des nombres naturels. Autrement dit, un ensemble est dénombrable lorsque ses éléments peuvent être listés sans omission ni répétition dans une suite indexée par les nombres naturels. Par exemple, l'ensemble des nombres entiers, l'ensemble des mots définis sur l'alphabet français et l'ensemble des suites composées de 0 et 1 sont tous les ensembles dénombrables.

En revanche, certains ensembles infinis, contiennent "trop" d'éléments pour être parcourus complètement par l'infinité des nombres naturels et sont donc "non dénombrables". Par exemple, l'ensemble des nombres réels est non dénombrable, car chaque nombre réel ne peut pas correspondre à un nombre naturel différent.

En informatique, les ensembles dénombrables sont souvent étudiés en terme d'**alphabet**, de **mot** et de **langage**.

Définition K.2.1

Un *alphabet* est un ensemble fini de caractères.

Par exemple, les 26 lettres en français constituent un alphabet. L'ensemble $\{0, 1\}$ est aussi un alphabet.

Définition K.2.2

Un mot défini sur un alphabet est une séquence finie composée par les éléments de cet alphabet.

Par exemple, “bonjour” est un mot défini sur l’alphabet des 26 lettres françaises. La séquence “010101010111111” est un mot défini sur l’alphabet $\{0, 1\}$.

Définition K.2.3

Un ensemble de mots définis sur un même alphabet est nommé “langage”.

Par exemple, le français est un langage défini sur l’alphabet des 26 lettres. L’ensemble $\{0, 1, 01, 11, 1010, 1111, 111110\}$ est un langage défini sur l’alphabet $\{0, 1\}$. Remarquons qu’un langage peut aussi bien être un ensemble fini qu’un ensemble infini. Tout dépend de ses règles de construction, nommées la grammaire. Plus précisément, un langage avec des longueurs de mots bornées est un ensemble fini.

L’exécution d’une machine de Turing sur un mot donné peut finir dans trois cas de figures :

1. Elle s’arrête dans un état **accepteur**.
2. Elle se trouve dans une configuration qui ne correspond à aucune transformation prévue. L’exécution finit donc dans un état non accepteur.
3. Elle ne s’arrête jamais.

Définition K.2.1

Un langage (ensemble) est “décidé” par une machine de Turing si pour tout mot du langage, l’exécution de la machine arrête dans un état accepteur. Et pour tout mot n’appartenant pas au langage, l’exécution finit dans un état non accepteur.

Définition K.2.2

Un langage (ensemble) est “accepté” par une machine de Turing si pour tout mot du langage, l’exécution de la machine arrête dans un état accepteur. Et pour tout mot n’appartenant pas au langage, l’exécution est infinie ou finit dans un état non accepteur.

Remarquons que si un langage est “décidé” par une machine de Turing, celle-ci donne une réponse définitive quel que soit le mot à traiter. En revanche, si un langage est seulement “accepté” par une machine de Turing, celle-ci risque de tomber dans un calcul infini quand il s’agit des mots n’appartenant pas au langage à reconnaître.

Selon ce principe nous pouvons diviser les ensembles dénombrables en trois

catégories : les ensembles récursifs, les ensembles récursivement énumérables, et les ensembles non récursivement énumérables.

- Un langage (ensemble) est dit “récursif” (désormais R) s’il est décidé par une machine de Turing. Par exemple, le langage L contenant tous les mots définis sur l’alphabet $\{a, b, c\}$ commençant par a est R, car on peut concevoir une machine de Turing qui, pour tout mot d’entrée w , lit le premier caractère et dit “oui” si w commence par un a et “non” dans le cas contraire. Cette machine accepte tous les mots du langage L et s’arrête sur tout mot d’entrée w .
- Un langage (ensemble) est récursivement énumérable (désormais RE) s’il est accepté par une machine de Turing. Un exemple connu des ensembles RE et non R est le problème d’arrêt des machine de Turing. Le langage défini par l’ensemble des programmes informatiques qui s’arrêtent est RE mais non R. Autrement dit, la machine de Turing qui dit si un programme donné s’arrête ou non ne donne pas toujours des réponses définitives.
- Un langage (ensemble) est non récursivement énumérable (désormais NRE) s’il est reconnu par aucune machine de Turing.

K.3 Fonction calculable

Selon Wolper (1991),

Définition K.3.1

Une machine de Turing calcule une fonction f , si pour tout mot d’entrée ω , elle s’arrête toujours dans une configuration où $f(\omega)$ se trouve sur le ruban.

Une fonction est dite “calculable” par une machine de Turing si elle est calculée par au moins une machine de Turing.

Selon la thèse de Turing-Church, toute procédure effective peut être simulée par une machine de Turing. Les fonctions calculables par les machines de Turing sont donc aussi nommées les “fonctions effectives”.

K.4 Ensemble de mesure nulle

La notion de l’ensemble de mesure nulle est indispensable pour la compréhension de suite aléatoire au sens de Martin-Löf, car ce dernier auteur utilise les ensembles de mesure nulle pour éliminer les suites régulières.

Une mesure est une fonction qui associe une “longueur”, un “volume” ou encore une “probabilité” à certaines parties d’un ensemble donné. Par exemple, la mesure de l’intervalle $[0, 0.5)$ dans l’ensemble $[0, 1)$ est de 0.5.

En termes mathématiques,

Définition K.4.1

Soit (χ, A) un espace mesurable (i.e. un couple (χ, A) où χ est un ensemble et A est une tribu sur χ). Une fonction μ définie sur A , à valeurs dans $[0, \infty]$ est appelée **mesure** lorsque les deux propriétés suivantes sont satisfaites :

- L’ensemble vide a une mesure nulle :

$$\mu(\emptyset) = 0$$

- L’application μ est σ -additive :
si E_1, E_2, \dots forment une famille dénombrable de parties de χ appartenant à A et si ces parties sont deux à deux disjointes, alors la mesure $\mu(E)$ de leur réunion E est égale à la somme des mesures des parties :

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k)$$

La mesure de Lebesgue est une manière très courant de définir la fonction calculant la mesure μ .

Définition K.4.2

Il existe une plus petite mesure définie sur une tribu de parties de l’espace \mathbb{R}^n qui soit complète et coïncide sur les pavés avec leur volume (c’est-à-dire le produit des longueurs de leurs côtés). Cette mesure est appelée la mesure de Lebesgue et sa tribu de définition la tribu de Lebesgue.

Par exemple, dans l’espace R , la mesure de Lebesgue de l’intervalle $[0, 1]$ est donnée par $\mu([0, 1]) = 1$. Fondé sur la notion de “mesure”, on peut définir :

Définition K.4.3

Soit (χ, A, μ) un espace mesuré. Une partie N de χ est dite négligeable (ou de mesure nulle) lorsqu’il existe un Y contenant N et de mesure nulle.

L. QUELQUES TECHNIQUES COMPLÉMENTAIRES

D'une manière générale, plus un texte est répétitif, plus il est compressible. Beaucoup d'algorithmes de compression sont donc précédés par des transformations réversibles qui ont pour objectif de regrouper les symboles identiques.

Le choix des techniques de transformation doit rester cohérent avec la méthode de compression utilisée. Les méthodes qui exploitent la fréquence locale des mots, notamment les codages d'entropie, ont souvent une meilleure performance après des transformations préalables. En revanche, les algorithmes fondés sur d'autres régularités pourraient être "pénalisés" par ces transformations, car certaines régularités structurelles peuvent disparaître suite à ces tentatives d'augmenter la fréquence locale des symboles.

Burrows-Wheeler-Transform La transformation de Burrows-Wheeler (BWT) ne compresses pas directement les données, elle se contente de les réorganiser de manière à obtenir des fréquences locales plus élevées des symboles. Elle "aide" les algorithmes de compression à obtenir de meilleurs résultats. Par exemple, le logiciel *zlib*¹ offre souvent un taux de compression plus élevé que le *gzip*², car le codage d'entropie dans le programme *zlib* est précédé par une BWT qui vise à augmenter les répétitions locales des fichiers.

Par exemple, la transformation BWT du message "TEXTE" se déroule en 2 étapes :

1. Tout d'abord, la chaîne de caractères est "rangée" dans un tableau carré en la décalant d'un caractère vers la droite à chaque nouvelle ligne (c.f tableau L.1).
2. Ensuite, on classe les différentes lignes par ordre alphabétique, comme il est montré dans le tableau L.2 :

Le texte après transformation est la dernière **colonne** du tableau L.2 précédée par son numéro de position 4, soit : "4TTXEE". En le comparant avec le texte d'origine

1. Disponible sur le site <http://www.zlib.net/>

2. Disponible sur le site <http://www.gzip.org>

POSITION	0	1	2	3	4
1	<i>T</i>	<i>E</i>	<i>X</i>	<i>T</i>	<i>E</i>
2	<i>E</i>	<i>T</i>	<i>E</i>	<i>X</i>	<i>T</i>
3	<i>T</i>	<i>E</i>	<i>T</i>	<i>E</i>	<i>X</i>
4	<i>X</i>	<i>T</i>	<i>E</i>	<i>T</i>	<i>E</i>
5	<i>E</i>	<i>X</i>	<i>T</i>	<i>E</i>	<i>T</i>

Tab. L.1: Début d'une transformation BWT.

POSITION	0	1	2	3	4
1(2)	<i>E</i>	<i>T</i>	<i>E</i>	<i>X</i>	<i>T</i>
2(5)	<i>E</i>	<i>X</i>	<i>T</i>	<i>E</i>	<i>T</i>
3(3)	<i>T</i>	<i>E</i>	<i>T</i>	<i>E</i>	<i>X</i>
4(1)	<i>T</i>	<i>E</i>	<i>X</i>	<i>T</i>	<i>E</i>
5(4)	<i>X</i>	<i>T</i>	<i>E</i>	<i>T</i>	<i>E</i>

Tab. L.2: Tableau de BWT après le classement alphabétique. Les numéros entre parenthèses désignent les positions des lignes dans le tableau L.1.

“TEXTE”, la transformation BWT ne réalise aucune compression immédiate. Cependant, le message obtenu présente sensiblement plus de répétitions locales. Et cette régularité sera exploitée par les algorithmes adaptés qui obtiendrait un meilleur taux de compression.

Pour que la transformation BWT soit sans perte, il faut que le texte d'origine puisse être reconstitué à partir du message “4TTXEE”. Le décodage d'une BWT est simple : il suffit de classer “4TTXEE” par ordre alphabétique, comme il est présenté dans le tableau L.3 : Le chiffre 4 au début du message codé indique que le

	0	1	2	3	4
Codé	<i>T</i>	<i>T</i>	<i>X</i>	<i>E</i>	<i>E</i>
Classé	<i>E</i>	<i>E</i>	<i>T</i>	<i>T</i>	<i>X</i>

Tab. L.3: Décodage d'une transformation BWT

décodage doit commencer par le 4ième caractère de la ligne “Classée” du tableau L.3. Le “T” qu'on observe à cette position est donc la première lettre du texte d'origine. Ce caractère étant le deuxième “T” de la ligne “Classée”, on recherche le deuxième “T” de la ligne “Codé” pour noter sa position, “2”. La deuxième position de la ligne “Classé” est occupée par un “E”, le deuxième caractère du texte d'origine est donc un “E”.

Ce “E” est le deuxième “E” de la ligne “Classé”, on retourne chercher le deuxième “E” de la ligne codée qui occupe la position 5. Comme la 5-ième position de la ligne “Classé” est un “X”, le troisième caractère est un “X”.

On répète l'opération jusqu'à la position 5 de la ligne "Codé" pour reconstituer le texte initial "TEXTE".

Move-To-Front Comme la BWT, la transformation move-to-front (désormais MTF) est aussi un traitement préparatoire des données. Les deux techniques se distinguent dans la mesure où BWT ne modifie jamais l'entropie d'un message donné, alors que MTF peut parfois réduire l'entropie d'un texte, surtout quand le message à compresser est localement répétitif à l'origine. Puisque la répétition locale peut être créée par BWT, la combinaison BWT + MTF peut parfois déboucher sur des résultats très satisfaisants.

Toujours avec un exemple concret, nous illustrons le principe de la MTF. Le message "textile" est transformé en 4 étapes :

1. Au départ, tous les symboles du message sont codés par ordre alphabétique. Dans notre exemple, le résultat de ce premier codage est présenté dans le tableau L.4 :

Symbole	<i>t</i>	<i>e</i>	<i>x</i>	<i>t</i>	<i>i</i>	<i>l</i>	<i>e</i>
Code	3	0	4	3	1	2	0

Tab. L.4: Tous les symboles sont codés par ordre alphabétique, 0 étant attribué à "e", "i" le suivant dans cet ordre recevra 1, etc ...

2. Ensuite, on transforme symbole par symbole les ordres alphabétiques obtenus dans le tableau L.4. Le premier ordre à traiter, "3", se trouve à la 4-ième position de la liste des ordres alphabétiques "01234". Le chiffre "4" sera donc retenu comme le début du texte final. En même temps, le quatrième chiffre du code initial ("3043120"), "3", est mis au début de la liste des ordres "01234". C'est-à-dire qu'à l'issue de la première étape, la liste des ordres devient "30124".
3. En consultant le tableau L.4, le deuxième symbole porte l'ordre "0", qui se trouve à la deuxième position de la liste des ordres actuels "30124". Le numéro "2" sera donc le deuxième symbole du texte final. Et le deuxième chiffre du code initial ("3043120"), "0", sera donc placé au début de la liste des ordres "30124", ce qui donne "03124".
4. En répétant l'opération jusqu'au dernier chiffre du tableau L.4, on obtient le code après transformation "4253455".

En comparant l'entropie du code initial "3043120" (cf l'équation L.1) et celle du code final "4253455" (cf l'équation L.2), on constate une baisse considérable de l'entropie de Shannon.

$$H(3043120) = \frac{2}{7} \log_2\left(\frac{7}{2}\right) + \frac{2}{7} \log_2\left(\frac{7}{2}\right) + 3 \times \frac{1}{7} \log_2 7 = 2.24 \quad (\text{L.1})$$

$$H(4253455) = \frac{2}{7} \log_2\left(\frac{7}{2}\right) + \frac{3}{7} \log_2\left(\frac{7}{3}\right) + 2 \times \frac{1}{7} \log_2 7 = 1.44 \quad (\text{L.2})$$

Ici, il est important de souligner que l'effet de la transformation MTF dépend fortement de la nature des données initiales. Plus les symboles sont localement concentrés dans un texte, plus la réduction d'entropie sera conséquente. Pour des données qui ne représentent aucune concentration locale, la transformation MTF pourrait même augmenter l'entropie de Shannon.

Bien que les techniques de transformation ne comprennent pas directement les données, elles facilitent incontestablement le travail des algorithmes de compression. Comme les algorithmes récents, notamment les codages par contexte, exploitent efficacement les régularités locales des données, l'utilisation de ces techniques est devenue quasiment systématique.

M. RÉSULTATS DE COMPRESSION OBTENUS SUR LES DONNÉES INTERNATIONALES

*Tab. M.1: Résultats des tests de compression effectués sur les 22 composantes de
CAC40, période d'observation 05/02/2001-09/02/2001*

	Titre	Taille	Taux1	Taux2	% zéro
1	Air France	8448	4,75%	0,75%	67,65%
2	Alcatel-Lucent	65536	9,25%	1,53%	76,90%
3	AXA	11520	5,81%	1,09%	66,79%
4	BNP Paribas	9472	5,58%	1,22%	62,34%
5	Bouygues	8448	5,69%	0,42%	65,42%
6	Cap Gemini	11520	6,07%	0,72%	64,58%
7	Carrefour	13056	6,51%	0,90%	65,79%
8	Danone	7168	5,94%	0,22%	64,77%
9	EADS	10496	2,75%	0,01%	70,23%
10	France Télécom	42240	12,34%	2,40%	67,07%
11	L'Oréal	10752	5,31%	0,30%	47,46%
12	Lafarge SA	10752	5,32%	0,57%	74,40%
13	Lagardère SCA	7427	5,71%	0,00%	65,19%
14	LVMH	10240	4,97%	0,24%	62,09%
15	Renault	7936	4,30%	-0,27%	70,30%
16	Sanofi-Aventis	7680	6,21%	0,01%	65,86%
17	Société générale	7680	3,32%	0,27%	61,68%
18	STMicronics	18432	10,22%	1,14%	57,73%
19	Suez eaux	7168	5,13%	0,73%	62,83%
20	Total	14848	7,10%	1,24%	65,51%
21	Veolia Environnement	8960	3,15%	-0,15%	63,86%
22	Vivendi	16896	7,90%	1,98%	64,46%
	Moyen		6,06%	0,70%	65,13%
	Ecart-type		2,26%	0,69%	5,75%

Annexe M

Tab. M.2: Tests de compression pour les 38 composantes de CAC40, période d'observation 24/11/2008-28/11/2008

Numéro	Titre	Taille	Taux 1	Taux 2	% zéro
1	Accor	21248	9,08%	0,11%	56,59%
2	Air France	18944	9,06%	0,75%	59,62%
3	Air liquide	24576	9,37%	0,02%	71,44%
4	Alcatel-Lucent	30720	10,43%	0,92%	61,79%
5	Alstom	62976	11,15%	0,74%	55,23%
6	Arcelor-Mittal	72704	13,44%	1,44%	57,95%
7	AXA	66816	12,82%	1,61%	63,99%
8	BNP Paribas	87552	12,22%	0,98%	54,94%
9	Bouygues	29952	9,25%	0,15%	57,55%
10	Cap Gemini	18944	8,59%	0,41%	54,80%
11	Carrefour	40704	10,50%	0,71%	55,99%
12	Crédit agricole	37888	12,03%	1,26%	64,26%
13	Danone	38656	9,96%	0,56%	54,12%
14	Eads	19200	8,85%	0,28%	61,08%
15	Edf	28416	8,21%	0,24%	54,20%
16	Essilor International	22528	6,38%	0,22%	59,66%
17	France Télécom	59904	12,01%	1,44%	61,48%
18	GDF	85504	13,39%	1,3%	56,42%
19	Oréal	42752	8,86%	0,33%	51,67%
20	Lafarge	53760	10,62%	0,45%	53,28%
21	Lagardère	10752	6,51%	-0,05%	53,10%
22	LVMH	43008	10,26%	0,35%	52,75%
23	Michelin	26368	9,63%	0,07%	50,90%
24	Pernod-Ricard	30976	7,53%	0,12%	58,38%
25	Peugeot	29952	11,27%	0,74%	58,57%
26	PPR	25856	9,74%	0,32%	50,32%
27	Renault	40192	11,24%	0,71%	56,62%
28	Saint-Gobain	39936	11,89%	0,77%	51,16%
29	Sanofi-Aventis	62976	11,31%	0,74%	55,60%
30	Schneider Electric	42240	9,75%	0,39%	52,70%
31	Société générale	73216	12,31%	0,92%	54,62%
32	STMicroelectronics	16384	7,80%	-0,06%	50,82%
33	Total	112384	14,53%	1,26%	56,59%
34	Unibail-Rodamco	17920	3,53%	-0,37%	55,75%
35	Vallourec	31232	7,63%	0,43%	45,15%
36	Veolia Environnement	25088	10,81%	0,74%	56,68%
37	Vinci	50176	11,53%	0,67%	57,61%
38	Vivendi	50432	11,62%	0,94%	61,04%
Moyen			10,16%	0,59%	56,35%
Ecart-type			2,23%	0,48%	4,75%

Tab. M.3: Tests de compression pour les 21 composantes de DJ, période d'observation
05/02/2001-09/02/2001

	Titre	Taille	Taux 1	Taux 2	% zéro
1	AXP (Amer Express Inc)	14592	16,91%	4,06%	39,34%
2	BMY(Bristol-Myers Squibb)	10752	10,03%	2,53%	37,55%
3	C (Citigroup)	19712	20,48%	5,17%	38,43%
4	DD (E.I. du Pont de Nem)	11008	18,01%	3,71%	36,10%
5	DIS (Walt Disney Company)	14848	18,68%	5,74%	40,24%
6	F (Ford Motor)	12032	17,25%	4,29%	42,27%
7	GE (Gen Electric Co)	33024	24,62%	7,11%	42,01%
8	HD (Home Depot)	20992	21,31%	4,77%	38,68%
9	IBM (Intl Business Mach)	23552	19,54%	3,51%	31,87%
10	JPM (JP Moorgan Co)	15360	16,60%	4,08%	42,31%
11	KO (Coca Cola)	11264	23,38%	3,81%	41,42%
12	MCD (Mcdonalds)	13059	17,88%	5,27%	45,16%
13	MO (Altria group)	16384	21,68%	6,49%	43,40%
14	MRK (Merck Co Inc)	12544	14,25%	3,13%	39,15%
15	PFE (Pfize Inc)	25600	21,78%	6,90%	40,92%
16	PG (Procter Gamble Co)	10240	11,62%	2,95%	42,87%
17	T (ATT Inc)	17152	21,71%	5,33%	44,11%
18	TXN (Texas Instruments)	34304	20,57%	5,39%	44,64%
19	VZ (Verizon Commun)	12544	18,33%	4,41%	41,06%
20	WMT (Wal-Mart)	18944	19,23%	5,34%	41,62%
21	XOM(Exxon Mobil Cp)	13056	16,50%	4,98%	40,95%
Moyen			18,59%	4,71%	40,67%
Ecart-type			3,60%	1,25%	3,08%

Annexe M

Tab. M.4: Tests de compression pour les 30 composantes de Nasdaq-100, période d'observation 05/02/2001-09/02/2001

	Titre	Taille	Taux 1	Taux 2	% zéro
1	Apple Inc. (AAPL)	31488	11,51%	3,44%	68,79%
2	Adobe Systems Incorporated (ADBE)	44032	18,94%	5,28%	57,17%
3	Akamai Technologies, Inc. (AKAM)	20480	12,84%	3,65%	57,31%
4	Altera Corporation (ALTR)	44544	14,86%	4,25%	61,20%
5	Amazon.com, Inc. (AMZN)	47104	13,41%	3,42%	66,02%
6	Amgen Inc. (AMGN)	47872	16,04%	5,00%	55,85%
7	Applied Materials, Inc. (AMAT)	111872	18,69%	5,32%	58,20%
8	Broadcom Corporation (BRCM)	260352	25,47%	6,18%	40,80%
9	Check Point Software Technologies Ltd. (CHKP)	97280	21,90%	4,98%	38,91%
10	Cisco Systems, Inc. (CSCO)	559872	23,65%	4,69%	58,86%
11	Citrix Systems, Inc. (CTXS)	30720	12,81%	4,18%	59,90%
12	Costco Wholesale Corporation (COST)	29184	13,64%	3,89%	62,46%
13	Dell Inc. (DELL)	118272	16,80%	4,26%	65,74%
14	EBay Inc. (EBAY)	38144	15,14%	4,58%	58,38%
15	Flextronics International Ltd. (FLEX)	52224	16,18%	4,49%	63,95%
16	Intel Corporation (INTC)	190720	18,55%	4,78%	61,15%
17	Intuit, Inc. (INTU)	34304	17,32%	4,59%	57,26%
18	KLA Tencor Corporation (KLAC)	62208	16,73%	4,33%	58,53%
19	Lam Research Corporation (LRCX)	19968	8,62%	2,47%	70,65%
20	Linear Technology Corporation (LLTC)	37376	14,38%	3,92%	55,61%
21	Maxim Integrated Products (MXIM)	37376	13,46%	3,80%	60,29%
22	Microsoft Corporation (MSFT)	180992	20,29%	5,13%	59,58%
23	NetApp, Inc. (NTAP)	174848	24,06%	6,14%	50,99%
24	Oracle Corporation (ORCL)	195840	20,23%	4,91%	60,61%
25	Paychex, Inc. (PAYX)	22016	12,55%	3,59%	59,70%
26	Qualcomm Incorporated (QCOM)	144128	23,38%	5,92%	49,17%
27	Starbucks Corporation (SBUX)	21760	11,51%	3,34%	66,29%
28	VeriSign, Inc. (VRSN)	72192	20,75%	4,96%	45,53%
29	Xilinx, Inc. (XLNX)	72192	18,83%	5,03%	53,84%
30	Yahoo! Inc. (YHOO)	95488	19,69%	5,00%	58,57%
Moyen			17,07%	4,57%	58,04%
Ecart-type			4,25%	0,83%	7,29%

N. ANALYSE EN COMPOSANTES PRINCIPALES (ACP)

Initiée par Pearson (1901), l'analyse en composantes principales (désormais ACP) est une méthode statistique qui révèle les principaux facteurs explicatifs de la divergence entre les objets observés. Selon le nombre des facteurs identifiés, nommés composantes principales, l'ACP peut éventuellement réduire le nombre de variables utilisées dans une base de données multi-variée.

Pour caractériser notre base de données à 4 variables (désormais "Base I"), l'ACP va identifier le nombre de facteurs nécessaires (composantes principales) pour expliquer la divergence entre les 111 actions observées. Ce nombre de composantes principales dépendra du degré de corrélation entre les 4 variables. Plus précisément, il est d'autant plus important que le degré d'indépendance entre les variables d'origine. Autrement dit, si les 4 variables dans Base I sont complètement indépendantes entre elles, l'ACP retiendra 4 composantes principales. En revanche, si la variable "Taux1" était parfaitement expliquée par "Taille", alors, le nombre de composantes principales retenues sera de 3, car "Taux1" ne fournit aucune information supplémentaire sur la différence entre les 111 actions observées.

Pour comprendre les relations entre les 4 variables utilisées, une ACP est conduite sur Base I selon les étapes suivantes :

- Dans un premier temps, Base I est représentée par une matrice 111×4 , nommée I , avec chaque ligne correspondant à une action observée et chaque colonne à l'une des 4 variables caractéristiques. A noter que chaque colonne de I est centrée et réduite pour que les résultats ne soient pas influencés par les échelles de mesure des différentes variables.

Pour vérifier la pertinence d'une ACP, l'indice KMO (Kaiser-Meyer-Olkin) et le test de Bartlett ont été réalisés sur Base I. L'indice KMO s'élevant à $51,11\% > 50\%$, le test de sphéricité de Bartlett délivrant une P-value de $6.83e - 05 < 0,05$, nous estimons que Base I est adéquate à l'ACP.

- Ensuite, la matrice de covariance de I , notée C , est calculée comme suit :

$$C = \begin{pmatrix} Cov_{taille,taux1} & Cov_{taille,taux2} & Cov_{taille,taille} & Cov_{taille,\%0} \\ Cov_{taux1,taux1} & Cov_{taux1,taux2} & Cov_{taux1,taille} & Cov_{taux1,\%0} \\ Cov_{taux2,taux1} & Cov_{taux2,taux2} & Cov_{taux2,taille} & Cov_{taux2,\%0} \\ Cov_{\%0,taux1} & Cov_{\%0,taux2} & Cov_{\%0,taille} & Cov_{\%0,\%0} \end{pmatrix}$$

Où $Cov_{taille,taux1}$ désigne la corrélation entre les variables “Taille” et “Taux1” dans Base I.

- Puis, les valeurs propres et les vecteurs propres sont respectivement calculés pour la matrice C . Chaque vecteur propre correspond à un facteur explicatif de I . Et la valeur propre de chaque vecteur obtenu indique l’importance du facteur correspondant.

Selon la définition des vecteurs propres, à l’issue de cette étape, les 111 titres initialement décrits par 4 variables inter-dépendantes, sont désormais caractérisés par 4 composantes parfaitement **indépendantes**. Selon la valeur propre de chaque vecteur propre, la composante représentée peut être retenue ou éliminée dans l’interprétation de Base I.

Le tableau N.1 présente les 4 vecteurs propres de la matrice C , classés selon leur pouvoir explicatif. Le tableau N.2 expose l’importance de chaque composante identifiée dans le tableau N.1. La ligne “Standard deviation” montre les racines carrées des valeurs propres. La ligne “Proportion of Variance” présente la part de la variance totale expliquée par chacune des composantes identifiées. Plus précisément, notées $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, les valeurs propres des 4 composantes, alors, la proportion de la variance expliquée par la composante 1 est calculée comme $\frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}$. La dernière ligne du tableau N.2 montre la proportion accumulée de la variance expliquée.

Tab. N.1: Les 4 vecteurs propres de C qui correspond chacun à un facteur explicatif de Base I.

	PC1	PC2	PC3	PC4
Taille	-0.3124	0.7932	-0.4954	-0.1667
Taux1	-0.6250	0.0213	0.1720	0.7612
Taux2	-0.5857	-0.0192	0.5415	-0.6028
%zéro	0.4108	0.6083	0.6571	0.1718

Dans ce tableau les composantes sont classées selon leur pouvoir explicatif (i.e leur valeur propre).

Pour calculer les “scores” de chaque composante principale, la Base I est projetée sur un espace vectoriel de 2 dimensions, dont le vecteur caractéristique, F , est constitué par les deux premières colonnes ($PC1$ et $PC2$) du tableau N.1.

Tab. N.2: L'importance des composantes identifiées

	PC1	PC2	PC3	PC4
Standard Deviation	1.5534	1.0093	0.7035	0.2706
Proportion of Variance	0.6033	0.2547	0.1237	0.0183
Cumulative Proportion	0.6033	0.8579	0.9817	1.0000

Les résultats de l'ACP sur la matrice I . La ligne "Standard deviation" montre les racines carrées des valeurs propres des composantes identifiées. La ligne "Proportion of Variance" donne la partie de la variance totale expliquée par chacune des composantes identifiées. La ligne "Cumulative Proportion" expose la proportion accumulée de la variance expliquée.

La matrice I projetée, notée I' , est calculée selon l'équation suivante :

$$I' = I \times F = I \times \begin{pmatrix} -0.3124 & 0.7932 \\ -0.6250 & 0.0213 \\ -0.5857 & -0.0192 \\ 0.4108 & 0.6083 \end{pmatrix}$$

I' ainsi obtenu est par définition, une matrice 111×2 , avec chaque ligne représentant une action observée. Chaque colonne de I' indique les "scores" obtenus par les actions sur l'une des deux composantes principales.

BIBLIOGRAPHIE

- ACKERT, L. F., ET B. F. SMITH (1993) : "Stock Price Volatility, Ordinary Dividends, and Other Cash Flows to Shareholders," *Journal of Finance*, 48(4), 1147–60.
- AKAIKE, H. (1973) : "Maximum likelihood identification of Gaussian autoregressive moving average models," *Biometrika*, 60(2), 255–265.
- ALEXANDER, S. (1961) : "Price Movements in Speculative Markets : Trends or Random Walk," *Industrial Management Review*, 1(2), 7–26.
- ALPERT, M., ET H. RAIFFA (1982) : *A Progress Report on the Training of Probability Assessors*, vol. 34. Cambridge University Press.
- ANDREASSEN, P., ET S. KRAUS (1990) : "Judgmental Extrapolation and the Saliency of Change," *Journal of Forecasting*, 9(4), 347–72.
- ARNOTT, R., ET P. BERNSTEIN (2002) : "What Risk Premium is 'Normal' ?," *Financial Analysts Journal*, 58(2), 64–85.
- AUCAMP, D. C. (1993) : "On the Extensive Number of Plays to Achieve Superior Performance with the Geometric Mean Strategy," *Management science*, 39(9), 1163–1172.
- AUMANN, R. (1976) : "Agreeing to Disagree," *Annals of Statistics*, 4(6), 1236–1239.
- AYDOGAN, K., ET G. G. BOOTH (1988) : "Are There Long Cycles in Common Stock Returns ?," *Southern Economic Journal*, 55(1), 141–149.
- BACHELIER, L. (1900) : "Théorie de la spéculation," Ph.D. thesis, Annales scientifiques de l'École Normale Supérieure : III-17, english version in "The Random Character of Stock Market Prices" (Cootner 1964, Cambridge, MA : MIT Press, 17-78).
- BAILLIE, R. (1996) : "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73(1), 5–59.
- BAILLIE, R., ET T. BOLLERSLEV (1996) : "Fractionally integrated generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 74(1), 3–30.
- BAKER, M., ET J. WURGLER (2000) : "The Equity Share in New Issues and Aggregate Stock Returns," *Journal of Finance*, 55(3), 2219–2257.

- BALDUZZI, P., G. BERTOLA, ET S. FORESI (1995) : “Asset Price Dynamics and Infrequent Feedback Trades,” *Journal of Finance*, 50(5), 1747–66.
- BALL, R. (1989) : “What Do We Know About Stock Market Efficiency?,” Working Paper 89-05, Rochester, Business - Managerial Economics Research Center.
- BANZ, R. W. (1981) : “The Relationship Between Market Value and Return of Common Stocks,” *Journal of Financial Economics*, 9(1), 3–18.
- BARABASI, A., ET T. VICSEK (1991) : “Multifractality of Self-affine Fractals,” *Physics Review A*, 44(4), 2730–2733.
- BARBERIS, N., A. SHLEIFER, ET R. VISHNY (1998) : “A model of investor sentiment1,” *Journal of Financial Economics*, 49(3), 307–343.
- BASU, S. (1977) : “Investment Performance of Common Stocks in Relation to Their Price-Earning Ratios :A Test of the Efficient Market,” *Journal of Finance*, 12(3), 129–156.
- BATIFOULIER, P., J. MERCHERS, ET D. URRUTIAGUER (2002) : “Risk Aversion Or Myopia ? Choices in Repeated Gambles and Retirement Investments,” *Revue de Philosophie Economique*, 2(6), 37–56.
- BEAVER, W. H. (1981) : “Market Efficiency,” *The Accounting Review*, 56(1), 23–37.
- BELAIRE-FRANCH, J., ET K. OPONG (2005) : “A Variance Ratio Test of the Behaviour of Some FTSE Equity Indices Using Ranks and Signs,” *Review of Quantitative Finance and Accounting*, 24(1), 93–107.
- BELGARD, C., ET P. GOLDSCHMIDT (1999) : “Forecasting Across Frequencies : Linearity and Nonlinearity,” Discussion paper, University of Western Australia Research Paper, Proceedings of the International conference on Advanced Technology.
- BEM, D. (1965) : “An Experimental Analysis of Self Persuasion,” *Journal of Experimental Social Psychology*, 1, 199–218.
- BERAN, J., ET Y. FENG (2001) : “Local Polynomial Estimation with a FARIMA-GARCH Error Process,” *Bernoulli*, 7(5), 733–750.
- BERNARD, V., ET J. THOMAS (1990) : “Evidence that stock prices do not fully reflect the implications of current earnings for future earnings,” *Journal of Accounting and Economics*, 13(4), 305–340.
- BERNSTEIN, P. L. (2007) : *Capital Ideas Evolving*. John Wiley & Sons.
- BHARDWAJ, G., ET N. R. SWANSON (2006) : “An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series,” *Journal of Econometrics*, 131(1-2), 539–578.
- BLACK, F. (1972) : “Capital Market Equilibrium with Restricted Borrowing,” *Journal of Business*, 45(3), 444–454.
- (1986) : “Noise,” *Journal of Finance*, 41(5), 529–543.

- BLACK, F., M. JENSEN, ET M. SCHOLES (1972) : *The Capital Asset Pricing Model : Some Empirical Tests in Studies in the Theory of Capital Markets*. Prageger New York.
- BLUME, M. (1970) : "Portfolio Theory : A Step Towards its Practical Application," *Journal of Business*, 43(2), 152–74.
- BLUME, M., ET I. FRIEND (1973) : "A New Look at the Capital Asset Pricing Model," *Journal of Finance*, 28(1), 19–33.
- BOEHME, R. D., ET S. M. SORESCU (2002) : "The Long-Run Performance Following Dividend Initiations and Resumptions : Underreaction or Product of Chance?," *Journal of Finance*, 57(2), 871–900.
- BOLLERSLEV, T. (1986) : "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31(3), 307–27.
- BOLLERSLEV, T., ET D. JUBINSKI (1999) : "Equity Trading Volume and Volatility : Latent Information Arrivals and Common Long-Run Dependencies," *Journal of Business & Economic Statistics*, 17(1), 9–21.
- BOUCHAUD, J. (2002) : "An Introduction to Statistical Finance," *Physica A*, 313, 238–251.
- BOUCHAUD, J., ET M. POTTERS (2000) : *Theory of Financial Risks From Statistical Physics to Risk Management*. Cambridge University Press.
- BOUDOUKH, J., M. RICHARDSON, ET T. SMITH (1993) : "Is the Ex ante Risk Premium Always Positive? A New Approach to Testing Conditional Asset Pricing Models," *Journal of Financial Economics*, 34(3), 387–408.
- BOURGHELLE, D., ET P. HYME (2009) : "De l'efficience des marchés au krach : l'illusion de la liquidité boursière," AFFI, Brest.
- BOX, G. E. P., ET D. A. PIERCE (1970) : "Distribution of the Autocorrelations in Autoregressive Moving Average Time Series Models," *Journal of American Statistical Association*, 65(332), 1509–1526.
- BOYNTON, W., ET H. R. OPPENHEIMER (2006) : "Anomalies in Stock Market Pricing : Problems in Return Measurements," *Journal of Business*, 79(5), 2617–2632.
- BRANDOUY, O., W. BRIEC, K. KERSTENS, ET I. V. DE WOESTYNE (2010) : "Portfolio Performance Gauging in Discrete Time Using a Luenberger Productivity Indicator," *Journal of Banking and Finance*, 34(8), 1899–1910.
- BRANDOUY, O., P. MATHIEU, ET I. VERYZHENKO (2009) : "Ex-Post Optimal Strategy for the Trading of a Single Financial Asset," *SSRN eLibrary*.
- BRAV, A., C. GECZY, ET P. A. GOMPERS (2000) : "Is the abnormal return following equity issuances anomalous?," *Journal of Financial Economics*, 56(2), 209–249.
- BRAV, A., ET J. HEATON (2002) : "Competing Theories of Financial Anomalies," *The Review of Financial Studies*, 15(2), 575–606.

- BREALEY, R. (1970) : "The Distribution and Independence of Successive Rates of Return from the British Equity Market," *Journal of Business Finance*, 2, 29–40.
- BROCK, W. A., J. A. SCHEINKMAN, W. D. DECHERT, ET B. LEBARON (1996) : "A test for independence based on the correlation dimension," *Econometric Reviews*, 15(3), 197–235.
- BROOKS, C. (1998) : "Chaos in Foreign Exchange Markets : A Sceptical View," *Computational Economics*, 11(3), 265–81.
- BURELL, O. (1951) : "Possibility of an Experimental Approach to Investment Studies," *Journal of Finance*, 6(2), 211–219.
- CALVET, L., ET A. FISHER (2002) : "Multifractality in Asset Returns : Theory and Evidence," *Review of Economics and Statistics*, 84(3), 381–406.
- CARHART, M. M. (1997) : "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52(1), 57–82.
- CHAITIN, G. (1969) : "On the length of programs for computing finite binary sequences : statistical considerations," *Journal of the ACM*, 16, 145–159.
- (1987) : *Algorithmic Information Theory*. Cambridge University Press, Cambridge Tracts in Theoretical Computer Science 1.
- CHAN, K. S., ET H. TONG (1986) : "On Estimating Thresholds in Autoregressive Models," *Journal of Time Series Analysis*, 7(3), 179–190.
- CHANG, E. C., ET W. G. LEWELLEN (1984) : "Market Timing and Mutual Fund Investment Performance," *Journal of Business*, 57(1), 57–72.
- CHEN, J., ET H. HONG (2002) : "Discussion of "Momentum and Autocorrelation in Stock Returns"," *The Review of Financial Studies*, 15(2), 565–573.
- CHEVALIER, J., ET G. ELLISON (1999) : "Are Some Mutual Fund Managers Better Than Others ? Cross-Sectional Patterns in Behavior and Performance," *Journal of Finance*, 54(3), 875–899.
- CHIARELLA, C., R. DIECI, ET X. HE (2007) : "Heterogeneous expectations and speculative behavior in a dynamic multi-asset framework," *Journal of Economic Behavior & Organization*, 62(3), 408–427.
- CHRISTENSEN, M. M. (2005) : "A Thesis on the Growth Optimal Portfolio and the Theory of Arbitrage Pricing and Portfolio Selection," Ph.D. thesis, University of Southern Denmark.
- CHUNG, Y. P., ET M. J. SCHILL (2006) : "Asset Pricing When Returns Are Non-normal : Fama-French Factors versus Higher-Order Systematic Comoments," *Journal of Business*, 79(2), 923–940.
- CHURCH, A. (1936) : "An Unsolvable Problem of Elementary Number Theory," *American Journal of Mathematics*, 58, 345–363.
- CLARK, R., ET W. T. ZIEMBA (1987) : "Playing the Turn-Of-The-Year Effect with Index Futures," *Operations Research*, 35(6), 799–813.

- CLEARY, J., ET I. WITTEN (1984) : "Data Compression Using Adaptive Coding and Partial String Marching," *IEEE Transactions on Communications*, 32(4), 396–402.
- CLEGG, R. (2005) : "A Practical Guide to Measuring the Hurst Parameter," *Dept. Of Mathematics, University of York*, pp. 3–14.
- CONRAD, J., ET G. KAUL (1998) : "An Anatomy of Trading Strategies," *Review of Financial Studies*, 11(3), 489–519.
- CONSTANTINIDES, G. M. (2002) : "Rational Asset Prices," *Journal of Finance*, 57(4), 1567–1591.
- CONT, R. (2001) : "Empirical Properties of Asset Returns : Stylized facts and Statical Issues," *Quantitative Finance*, 1(2), 223–36.
- COOPER, M. J., R. C. GUTIERREZ, ET A. HAMEED (2004) : "Market States and Momentum," *Journal of Finance*, 59(3), 1345–1365.
- COWLES, A. (1933) : "Can Stock Market Forecasters Forecasts?," *Econometrica*, 1(2), 309–324.
- CROSS, F. (1973) : "The Behavior of Stock Prices on Fridays and Mondays," *Financial Analysts Journal*, 29(6), 67–69.
- CUNNINGHAM, S. W. (1973) : "The Predictability of British Stock Market Prices," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 22(3), 315–331.
- DACOROGNA, M. M., ET AL. (2001) : *An Introduction to High-Frequency Finance*. Academic Press, New York.
- DACOROGNA, M. M., U. A. MULLER, O. V. PICTET, ET C. D. VRIES (1995) : "The Distribution of Extremal Foreign Exchange Rate Returns in Extremely Large Data Sets," Working papers, Olsen and Associates.
- DANIEL, KENT, E. A. (1997) : "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks," *Journal of Finance*, 52(3), 1035–58.
- DANIEL, K., D. HIRSHLEIFER, ET A. SUBRAHMANYAM (1998) : "Investor Psychology and Security Market Under- and Overreactions," *Journal of Finance*, 53(6), 1839–1885.
- DAVIS, D. (1952) : "An Analysis of Some Failure Data," *Journal of American Statistical Association*, 47(258), 113–150.
- DEBONDT, W. (2002) : "Discussion of "Competing Theories of Financial Anomalies"," *The Review of Financial Studies*, 15(2), 607–613.
- DEBONDT, W., ET T. RICHARD (1985) : "Does the Stock Market Overreact?," *Journal of Finance*, 40(3), 793–805.
- DELAHAYE, J.-P. (1994) : *Information, Complexité et Hasard*. HERMES.
- DELAHAYE, J.-P. (2009a) : "Escroquerie ou jeu risqué?," *Pour la Science*, Novembre(385), 136–141.

- DELAHAYE, J.-P. (2009b) : “Presque tout est indécidable!,” *Pour la Science*, Février(375), 88–93.
- DELONG, SHLEIFER, SUMMERS, ET WALDMANN (1990) : “Noise Trader Risk in Financial Markets,” *Journal of Political Economy*, 98(4), 703–738.
- DERMAN, E. (2002) : *My Life as a Quant : Reflections on Physics and Finance*. John Wiley & Sons, United States.
- DERVEEUW, J., B. BEAUFILS, O. BRANDOUY, ET P. MATHIEU (2007) : “Testing Double Auction as a Component within a Generic Market Model Architecture,” *Lecture Notes in Economic and Mathematical Systems*, Spring, 47–61.
- DICKEY, D. A., ET W. A. FULLER (1979) : “Distribution of the Estimators for Autoregressive Time Series With a Unit Root,” *Journal of the American Statistical Association*, 74(366), 427–431.
- DIEBOLD, F. X., ET J. A. NASON (1990) : “Nonparametric exchange rate prediction?,” *Journal of International Economics*, 28(3-4), 315–332.
- DING, Z., C. GRANGER, ET R. ENGLE (1993) : “A Long Memory Property of Stock Market Returns and a New Model,” *Journal of Empirical Finance*, 1(1), 83–106.
- D’MELLO, R., S. P. FERRIS, ET C. Y. HWANG (2003) : “The tax-loss selling hypothesis, market liquidity, and price pressure around the turn-of-the-year,” *Journal of Financial Markets*, 6(1), 73–98.
- DOWNEY, R., ET D. HIRSCHFELDT (2010) : *Algorithmic Randomness and Complexity (Theory and Applications of Computability)*. Springer.
- DUNIS, C. L., ET M. WILLIAMS (2003) : *Applied Quantitative Methods for Trading and Investment* chap. Applications of Advanced Regression Analysis for trading and Investment, pp. 1–40. John Wiley & Sons.
- DUNN, S. (2001) : “Bounded Rationality Is Not Fundamental Uncertainty : A Post Keynesian Perspective,” *Journal of Post Keynesian Economics*, 23(4), 567–587.
- EDWARDS, W. (1968) : “Conservatism in Human Information Processing,” in *Formal Representation of Human Judgement*, ed. by B. Kleinmütz. Wiley, New York.
- ELESWARAPU, V., ET R. THOMPSON (2007) : “Testing for Negative Expected Market Return Premia,” *Journal of Banking and Finance*, 31(6), 1755–1770.
- ELTON, E. J. (1999) : “Presidential Address : Expected Return, Realized Return, and Asset Pricing Tests,” *Journal of Finance*, 54(4), 1199–1220.
- ELTON, E. J., M. GRUBER, S. DAS, ET M. HLAVKA (1993) : “Efficiency with Costly Information : A Reinterpretation of Evidence from Managed Portfolios,” *Review of Financial Studies*, 6(1), 1–22.
- ENGLE, R. (1982) : “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, 50(4), 987–1007.

- EPSTEIN, B., ET M. SOBEL (1953) : "Life Testing," *Journal of American Statistical Association*, 48(263), 486–502.
- FAMA, E. (1965) : "Behavior of Stock Market Prices," *Journal of Business*, 38(1), 34–105.
- (1970) : "Efficient Capital Markets : A Review of Theory and Empirical Work," *Journal of Finance*, 25(2), 383–417.
- (1976) : "Efficient Capital Markets : Reply," *Journal of Finance*, 31(1), 143–145.
- (1990) : "Stock Returns, Expected Returns, and Real Activity," *Journal of Finance*, 45(4), 1089–1108.
- (1991) : "Efficient Capital Markets : II," *Journal of Finance*, 46(5), 1575–1617.
- (1998) : "Market Efficiency, Long-Term Returns, and Behavioral Finance," *Journal of Financial Economics*, 49(3), 283–306.
- FAMA, E., ET K. FRENCH (1988) : "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, 96(2), 246–73.
- (1992) : "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47(2), 427–465.
- (1993) : "Common Risk Factors in the Returns on Stock and Bonds," *Journal of Financial Economics*, 33(1), 3–56.
- (1996) : "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance*, 51(1), 55–84.
- (2000) : "Forecasting Profitability and Earnings," *Journal of Business*, 73(2), 161–75.
- (2004) : "The Capital Asset Pricing Model : Theory and Evidence," *Journal of Economic Perspectives*, 18(3), 25–46.
- FAMA, E., ET J. MACBETH (1973) : "Risk, Return, and Equilibrium : Empirical Tests," *Journal of Political Economy*, 81(3), 607–36.
- FRENCH, K. (1980) : "Stock Returns and the Weekend Effect," *Journal of Financial Economics*, 8(1), 55–69.
- GEWEKE, J., ET S. PORTER-HUDAK (1983) : "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4(4), 221–238.
- GIBBONS, M., S. ROSS, ET J. SHANKEN (1989) : "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57(5), 1121–152.
- GIBBONS, M. R., ET P. HESS (1981) : "Day of the Week Effects and Asset Returns," *Journal of Business*, 54(4), 579–96.
- GIRAITIS, L., P. KOKOSZKA, ET R. LEIPUS (2001) : "Testing for Long Memory in the Presence of a General Trend," *Journal of Applied Probability*, 38(4), 1033–1054.

- GLOSTEN, L. R., ET P. R. MILGROM (1985) : “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders,” *Journal of Financial Economics*, 14(1), 71–100.
- GÖDEL, K. (1931) : “Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. (Sur quelques théorèmes formels indécidables des Principia Mathematica et des systèmes qui leur sont reliés),” *Monatshefte für Mathematik und Physik*, 38, 173–198, Traduit en anglais par van Heijenoort in *From Frege to Gödel*. Harvard University Press, 1971. pp. 596-616.
- GOETZMANN, W. N., ET M. MASSA (2002) : “Daily Momentum and Contrarian Behavior of Index Fund Investors,” *Journal of Financial and Quantitative Analysis*, 37(3), 375–389.
- GRABBE, J. (1996) : *International Financial Markets*. Prentice Hall, 3rd edn.
- GRANGER, C. (1966) : “The Typical Spectral Shape of an Economic Variable,” *Econometrica*, 34(1), 150–161.
- (1986) : “Developments in the Study of Cointegrated Economic Variables,” *Oxford Bulletin of Economics and Statistics*, 48(3), 213–28.
- GRANGER, C., ET Z. DING (1994) : “Stylized Facts on the Temporal and Distributional Properties of Daily Data from Speculative Markets,” University of California at San Diego, Economics Working Paper Series 94-19, Department of Economics, UC San Diego.
- GRANGER, C., ET N. HYUNG (2004) : “Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns,” *Journal of Empirical Finance*, 11(3), 399–421.
- GRANGER, C., ET O. MORGENSTERN (1963) : “Spectral Analysis of New York Stock Market Prices,” *Kyklos*, 16, 1–27, Reprinted in *The Random Character of Stock Market Prices*, edited by P.H. Cootner, The MIT Press, 1964.
- GREENE, M. T., ET B. D. FIELITZ (1977) : “Long-term dependence in common stock returns,” *Journal of Financial Economics*, 4(3), 339–349.
- GRINBLATT, M., R. MASULIS, ET S. TITMAN (1984) : “The valuation effects of stock splits and stock dividends,” *Journal of Financial Economics*, 13(4), 461–490.
- GRINBLATT, M., ET S. TITMAN (1992) : “The Persistence of Mutual Fund Performance,” *Journal of Finance*, 47(5), 1977–84.
- GRINOLD, R. C., ET R. N. KAHN (1999) : *Active Portfolio Management*. MacGraw-Hill, 2nd edn.
- GROSSMAN, S., ET J. STIGLITZ (1980) : “On the Impossibility of Informationally Efficient Markets,” *American Economic Review*, 70(3), 393–408.
- HALL, P. (1982) : “On Some Simple Estimates of an Exponent of Regular Variation,” *Journal of the Royal Statistical Society : Series B*, 44, 37–42.
- HAMILTON, J. (1989) : “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle,” *Econometrica*, 57(2), 357–84.

- (1994) : *Time Series Analysis*. Princeton University Press.
- HARRIS, L. (1990) : “Estimation of Stock Price Variances and Serial Covariances from Discrete Observations,” *Journal of Financial and Quantitative Analysis*, 25(3), 291–306.
- HARVEY, C. R., ET A. SIDDIQUE (2000) : “Conditional Skewness in Asset Pricing Tests,” *Journal of Finance*, 55(3), 1263–1295.
- HAUGEN, R., ET J. LAKONISHOK (1988) : *The Incredible January Effect*. Homewood : Dow Jones-Irwin.
- HELLSTRÖM, T., ET K. HOLMSTRÖM (1998) : “Predicting the Stock Market,” Discussion paper, Department of Mathematics and Physics, Mälardalen University, Sweden.
- HILL, B. (1975) : “A Simple General Approche to Inference about the Tail of Distribution,” *Annals of Statistics*, 3(5), 1163–1174.
- HONG, H., T. LIM, ET J. C. STEIN (2000) : “Bad News Travels Slowly : Size, Analyst Coverage, and the Profitability of Momentum Strategies,” *Journal of Finance*, 55(1), 265–295.
- HONG, H., ET J. C. STEIN (1999) : “A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets,” *Journal of Finance*, 54(6), 2143–2184.
- HSIEH, D. A. (1991) : “Chaos and Nonlinear Dynamics : Application to Financial Markets,” *Journal of Finance*, 46(5), 1839–77.
- (1995) : “Nonlinear Dynamics in Financial Markets : Evidence and Implications,” *Financial Analysts Journal*, 51(4), 55–62.
- HUFFMAN, D. (1952) : “A Method for the Construction of Minimum-Redundancy Codes,” *Proceedings of the Institute of Radio Engineers*, 40, 1098–1101.
- HURST, H. (1951) : “Long Term Storage Capacity of Reservoirs,” *Transactions of the American Society of Civil Engineers*, 116, 770–799.
- IKENBERRY, D. L., G. RANKINE, ET E. K. STICE (1996) : “What Do Stock Splits Really Signal?,” *Journal of Financial and Quantitative Analysis*, 31(3), 357–375.
- IPPOLITO, R. A. (1989) : “Efficiency with Costly Information : A Study of Mutual Fund Performance, 1965-1984,” *The Quarterly Journal of Economics*, 104(1), 1–23.
- JAFFE, J. F., ET R. WESTERFIELD (1985) : “The Week-End Effect in Common Stock Returns : The International Evidence,” *Journal of Finance*, 40(2), 433–54.
- JANSEN, D., ET C. VRIES (1991) : “On the Frequency of Large Stock Returns : Putting Booms and Busts into Perspective,” *Review of Economics and Statistics*, 73(1), 18–24.

- JAWADI, F., ET Y. KOUBBAA (2007) : “Dynamique non-linéaire des marchés boursiers du G7 : Une application des modèles STAR,” *Finance*, 28(1), 2007.
- JEGADEESH, N. (2001) : “Profitability of Momentum Strategies : An Evaluation of Alternative Explanations,” *Journal of Finance*, 56(2), 699–720.
- JEGADEESH, N., ET S. TITMAN (1993) : “Returns to Buying Winners and Selling Losers : Implications for Stock Market Efficiency,” *Journal of Finance*, 48(1), 65–91.
- (2002) : “Cross-Sectional and Time-Series Determinants of Momentum Returns,” *The Review of Financial Studies*, 15(1), 143–157.
- JENSEN, M. (1968) : “The Performance of Mutual Funds in the Period 1945-1964,” *Journal of Finance*, 32(2), 389–416.
- JOHNSON, T. C. (2002) : “Rational Momentum Effects,” *Journal of Finance*, 57(2), 585–608.
- KAHNEMAN, D. (2003) : “Maps of Bounded Rationality : Psychology for Behavioral Economics,” *American Economic Review*, 93(5), 1449–1475.
- KAHNEMAN, D., ET A. TVERSKY (1979) : “Prospect theory : An analysis of decision under risk,” *Econometrica*, 47(2), 263–291.
- KAMINSKI, K., ET A. LO (2007) : “When Do Stop-Loss Rules Stop Losses?,” Discussion paper, MIT Operation research Center.
- KEIM, D. B. (1983) : “Size-related anomalies and stock return seasonality : Further empirical evidence,” *Journal of Financial Economics*, 12(1), 13–32.
- KEIM, D. B., ET R. F. STAMBAUGH (1984) : “A Further Investigation of the Weekend Effect in Stock Returns,” *Journal of Finance*, 39(3), 819–35.
- KENDALL, M., ET A. B. HILL (1953) : “The Analysis of Economic Time-Series-Part I : Prices,” *Journal of the Royal Statistical Society*, 116(1), 11–34.
- KEYNES, J. (1936) : *The General Theory of Employment, Interest and Money*. Macmillan Cambridge University Press.
- KEYNES, J. M. (1921) : *A Treatise on Probability*. London : Macmillan.
- KING, W. (1930) : *Index Numbers Elucidated*.
- KLEIDON, A. W. (1986) : “Variance Bounds Tests and Stock Price Valuation Models,” *Journal of Political Economy*, 94(5), 953–1001.
- KNIGHT, F. (1921) : *Risk, uncertainty and Profit*. Boston and New York : Houghton Mifflin, Hart, Schaffner, and Marx Prize Essays, no. 31.
- KOLMOGOROV, A. (1965) : “Three Approaches to the Quantitative Definition of Information,” *Problems of Information Transmission*, 1(1), 1–11.
- KORAJCZYK, R., ET R. SADKA (2004) : “Are Momentum Profits Robust to Trading Costs?,” *Journal of Finance*, 59(3), 1039–1082.
- KOSOWSKI, R., A. TIMMERMANN, R. WERMERS, ET H. WHITE (2006) : “Can Mutual Fund “Stars” Really Pick Stocks? New Evidence from a Bootstrap Analysis,” *Journal of Finance*, 61(6), 2551–2595.

- KOTHARI, S., ET J. SHANKEN (1997) : “Book-to-Market, Dividend Yield, and Expected Market Returns : a Time-Series,” *Journal of Financial Economics*, 44(3), 169–203.
- KOTHARI, S., J. SHANKEN, ET R. SLOAN (1995) : “Another Look at the Cross-section of Expected Returns,” *Journal of Finance*, 50(1), 185–224.
- KWIATKOWSKI, D., P. C. B. PHILLIPS, P. SCHMIDT, ET Y. SHIN (1992) : “Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root?,” *Journal of Econometrics*, 54(1-3), 159–178.
- L. HAAN, D.W. JANSENAND, K. K., ET C. VRIES (1991) : *Safety First Portfolio Selection : Extreme Value theory and Long Run Asset Risks*, vol. Extreme Value Theory and Applications, Boston. J. Calambos and al. (eds.).
- LAKONISHOK, J., A. SHLEIFER, ET R. VISHNY (1994) : “Contrarian Investment, Extrapolation, and Risk,” *Journal of Finance*, 49(5), 1541–1578.
- LAKONISHOK, J., ET T. VERMAELEN (1990) : “Anomalous Price Behavior around Repurchase Tender Offers,” *Journal of Finance*, 45(2), 455–77.
- LANE, J. A., D. A. PEEL, ET E. J. RAEBURN (1996) : “Some Empirical Evidence on the Time-Series Properties of Four UK Asset Prices,” *Economica*, 63(251), 405–426.
- LAPORTA, R., J. LAKONISHOK, A. SHLEIFER, ET R. VISHNY (1997) : “Good News for Value Stocks : Further Evidence on Market Efficiency,” *Journal of Finance*, 52(2), 859–74.
- LATHAM, M. (1986) : “Information efficiency and information subsets,” *Journal of Finance*, 41(1), 39–52.
- LEE, C.-F., ET A. LEE (2006) : *Encyclopedia of Finance*. Springer Science, New York.
- LEFÈVRE, E. (1923) : *Reminiscences of a Stock Operator*. The Sun Dial Press.
- LEROY, S. (1989) : “Efficient Capital Markets and Martingales,” *Journal of Economic Literature*, 27(4), 1583–1621.
- (2004) : “Rational Exuberance,” *Journal of Finance*, 42(3), 783–804.
- LEROY, S., ET R. PORTER (1981) : “The Present-Value Relation : Tests Based on Implied Variance Bounds,” *Econometrica*, 49(3), 555–74.
- LEVY, P. (1937) : *Théorie de l'addition des variables aléatoires*. Gauthier-Villars.
- LEWELLEN, J. (2002) : “Momentum and Autocorrelation in Stock Returns,” *The Review of Financial Studies*, 15(2), 533–563.
- LEWIS, D. (1969) : *Convention : A Philosophical Study*. Harvard University Press, Cambridge.
- LIANG, B. (2001) : “On the Performance of Hedge Funds,” *Financial Analysts Journal*, 55(4), 72–85.

- LINTNER, J. (1965) : “The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets,” *Review of Economics and Statistics*, 47(2), 13–37.
- LJUNG, G., ET G. BOX (1978) : “On a Measure of Lack of Fit in Time Series Models,” *Biometrika*, 66(2), 297–303.
- LO, A. (1991) : “Long-Term Memory in Stock Market Prices,” *Econometrica*, 59(5), 1279–1313.
- LO, A., ET A. MACKINLAY (1988) : “Stock Market Prices do not Follow Random Walks : Evidence from a Simple Specification Test,” *Review of Financial Studies*, 1(1), 41–66.
- (1990) : “Data-Snooping Biases in Tests of Financial Asset Pricing Models,” *Review of Financial Studies*, 3(3), 431–468.
- (2002) : *A Non-Random Walk Down Wall Street*. Princeton University Press, The United Kingdom.
- LOBATO, I., J. C. NANKERVIS, ET N. E. SAVIN (2001) : “Testing for Autocorrelation Using a Modified Box-Pierce Q Test,” *International Economic Review*, 42(1), 187–205.
- LOBATO, I. N., ET P. M. ROBINSON (1998) : “A Nonparametric Test for $I(0)$,” *Review of Economic Studies*, 65(3), 475–95.
- LOBATO, I. N., ET N. E. SAVIN (1998) : “Real and Spurious Long-Memory Properties of Stock-Market Data,” *Journal of Business and Economic Statistics*, 16(3), 261–268.
- LORETAN, M., ET P. PHILIPPS (1994) : “Testing the Covariance Stationarity of Heavy Tailed Time Series : An overview of the Theory with Applications to Several Financial Datasets,” *Journal of Empirical Finance*, 1(2), 18–24.
- LOUGHRAN, T., ET J. R. RITTER (1996) : “Long-Term Market Overreaction : The Effect of Low-Priced Stocks,” *Journal of Finance*, 51(5), 1959–70.
- LUCAS, R. (1978) : “Asset Prices in an Exchange Economy,” *Econometrica*, 46(4), 1429–45.
- (1980) : “Methods and Problems in Business Cycle Theory,” *Journal of Money, Credit and Banking*, 12(4), 696–715.
- LUX, T., ET S. SCHORNSTEIN (2005) : “Genetic learning as an explanation of stylized facts of foreign exchange markets,” *Journal of Mathematical Economics*, 41(1-2), 169–196.
- MA, L. (2009) : “Are Stylized Facts Quantitatively Universal?,” *AFFI 2009, Brest*.
- MALKIEL, B. (2003) : “The Efficient Market Hypothesis and its Critiques,” *Journal of Economic Perspectives*, 17(1), 59–82.
- MANDELBROT, B. (1963) : “The Variation of Certain Speculative Prices,” *Journal of Business*, 36(4), 394–419.

- (1965) : “Une Classe de Processus Stochastiques Homothétiques à Soi : Application à la loi Climatologique de H.E.Hurst,” *Comptes Rendus (Paris)*, 260, 3274–7.
- MANDELBROT, B. (1997) : *Fractal, Hasard et Finance*. Flammarion.
- MANDELBROT, B. B. (1971) : “When Can Price Be Arbitraged Efficiently? A Limit to the Validity of the Random Walk and Martingale Models,” *The Review of Economics and Statistics*, 53(3), 225–36.
- MANSILIA, R. (2004) : “Algorithmic Complexity in Real Financial Markets,” Unpublished Working Paper of the Center for Interdisciplinary Research in Sciences and the Humanities of the National University of Mexico.
- MARKOWITZ, H. (1959) : *Portfolio Selection Efficient Diversification of Investments*. John Wiley & Sons.
- MARTIKAINEN, T., ET V. PUTTONEN (1996) : “Finish Day-of-the-week Effects,” *Journal of Business Finance & Accounting*, 23(7), 1019–1032.
- MARTIN-LÖF, P. (1966) : “The Definition of Random Sequences,” *Information and Control*, 9, 113–150.
- MASON, D., ET T. TUROVA (1994) : *Weak Convergence of the Hill Estimator Process, In Extreme Value Theory and Applications*. Kluwer, Dordrecht.
- MATTEO, T. D. (2007) : “Multi-scaling in Finance,” *Quantitative Finance*, 7(1), 21–36.
- MCCAULEY, J. (2006) : “Response to Worrying Trends in Econophysics,” *Munich Personal RePEc Archive*, Department of Economics University of Houston.
- MICHAELY, R., R. THALER, ET K. WOMACK (1995) : “Price Reactions to Dividend Initiations and Omissions : Overreaction or Drift?,” *Journal of Finance*, 50(2), 573–608.
- MIGNON, V. (1998) : *Marchés financiers et modélisation des rentabilités financières*. Economica.
- MILGROM, P., ET N. STOKEY (1982) : “Information, trade and common knowledge,” *Journal of Economic Theory*, 26(1), 17–27.
- MILLER, M. H. (1986) : “Behavioral Rationality in Finance : The Case of Dividends,” *Journal of Business*, 59(4), 451–68.
- MILLS, T. (1999) : *The Econometric Modeling of Financial Time Series*. Cambridge University Press, 2nd edn.
- MÜLLER, U., M. M. DACOROGNA, R. B. OLSEN, O. V. PICTET, M. SCHWARTZ, ET C. MORGENEGG (1990) : “Statistical Study of Foreign Exchange Rates, Empirical Evidence of a Price Change Scaling Law and Intraday Analysis,” *Journal of Banking and Finance*, 14(6), 1189–1208.
- MUTH, J. (1961) : “Rational Expectations and the Theory of Price Movements,” *Econometrica*, 29(6), 315–335.

- NELSON, D. B. (1991) : “Conditional Heteroskedasticity in Asset Returns : A New Approach,” *Econometrica*, 59(2), 347–370.
- ORLÉAN, A. (1989) : “Pour une approche cognitive des conventions économiques,” *Revue économique*, 40(2), 241–272.
- (1999) : *Le pouvoir de la finance*. Odile Jacob.
- OSBORNE, M. (1959) : “Brownian Motion in the Stock Market,” *Operations Research*, 7(2), 145–173.
- OSTDIEK, B. (1998) : “The World Ex Ante Risk Premium : an Empirical Investigation,” *Journal of International Money and Finance*, 17(1), 967–999.
- PEARSON, K. (1901) : “On Lines and Planes of Closest Fit to Systems of Points in Space,” *Philosophical Magazine*, 2(6), 559–572.
- PETRUCCELLI, J. D., ET S. W. WOOLFORD (1984) : “A Threshold AR(1) Model,” *Journal of Applied Probability*, 21(2), 270–286.
- PHILLIPS, P. C. B., ET P. PERRON (1988) : “Testing for a Unit Root in Time Series Regression,” *Biometrika*, 75(2), 335–346.
- POITRAS, G. (2002) : “The Philosophy of Investment : a Post Keynesian Perspective,” *Journal of Post Keynesian Economics*, 25(1), 105–121.
- POTTER, M., ET J.-P. BOUCHAUD (1997) : *Théorie des Risques Financiers*. CEA.
- R DEVELOPMENT CORE TEAM (2005) : *R : A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
- RAY, B., ET R. TSAY (2000) : “Long-range Dependence in Daily Stock Volatilities,” *Journal of Business & Economic Statistics*, 18(2), 254–262.
- REINGANUM, M. R. (1983) : “The anomalous stock market behavior of small firms in January : Empirical tests for tax-loss selling effects,” *Journal of Financial Economics*, 12(1), 89–104.
- RICK, S., ET G. LOEWENSTEIN (2008) : *Handbook of Emotions* chap. 9 The Role of Emotion in Economic Behavior, pp. 138–152. The Guilford Press, 3th edn.
- RITTER, J. R. (1988) : “The Buying and Selling Behavior of Individual Investors at the Turn of the Year,” *Journal of Finance*, 43(3), 701–17.
- ROBERTS, H. V. (1959) : “Stock-Market "Patterns" and Financial Analysis : Methodological Suggestions,” *Journal of Finance*, 14(1), 1–10.
- ROBINSON, P. M. (1994) : “Semiparametric Analysis of Long-Memory Time Series,” *The Annals of Statistics*, 22(1), 515–539.
- ROLL, R. (1977) : “A Critique of the Asset Pricing Theory’s Tests’ Part I : On Past and Potential Testability of the Theory,” *Journal of Financial Economics*, 4(2), 129–176.
- (1983) : “Vas ist das ? The turn of the year effect and the return premium of small firms,” *Journal of Portfolio Management*, 9(2), 18–28.

- (1984) : “A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market,” *American Financial Association*, 39(4), 1127–1139.
- ROSS, S. A. (1976) : “The Arbitrage Theory of Capital Asset Pricing,” *Journal of Economic Theory*, 13(3), 341–60.
- (1999) : “Adding Risks : Samuelson’s Fallacy of Large Numbers Revisited,” *Journal of Financial and Quantitative Analysis*, 34(03), 323–339.
- (2002) : “Neoclassical Finance, Alternative Finance and the Closed End Fund Puzzle,” *European Financial Management*, 8(2), 129–137.
- (2004) : *Neoclassical Finance*. Princeton University Press.
- ROUWENHORST, K. (1998) : “International Momentum Strategies,” *Journal of Finance*, 53(1), 267–284.
- ROZEFF, M., ET W. KINNEY (1976) : “Capital Market Seasonality : The Case of Stock Returns,” *Journal of Financial Economics*, 3(4), 379–402.
- RUBINSTEIN, M. (1975) : “Securities Market Efficiency in an Arrow-Debreu Economy,” *American Economic Review*, 65(5), 812–824.
- (2001) : “Rational Markets : Yes or No ? The Affirmative Case,” *Financial Analysts Journal*, 57(3), 15–29.
- SAMUELSON, P. (1965) : “Proof that Properly Anticipated Prices Fluctuate Randomly,” *Industrial Management Review*, 6(2), 41–49.
- SARANTIS, N. (2001) : “Nonlinearities, cyclical behaviour and predictability in stock markets : international evidence,” *International Journal of Forecasting*, 17(3), 459–482.
- SAVAGE, L. (1954) : *The foundations of statistics*. New York : John Wiley & Sons.
- SCHEINKMAN, J. A., ET B. LEBARON (1989) : “Nonlinear Dynamics and Stock Returns,” *Journal of Business*, 62(3), 311–37.
- SCHELLING, T. (1960) : *The strategic of conflict*. Harvard University Press.
- SCHNORR, C. P. (1971) : “An Unified Approach to the Definition of a Random Sequence,” *Mathematical Systems Theory*, 5(3), 246–258.
- SCHWARZ, G. (1978) : “Estimating the Dimension of a Model,” *The Annals of Statistics*, 6(2), 461–464.
- SHANNON, C. (1948) : “A mathematical theory of communication,” *Bell System Technical Journal*, 27, 379–423.
- SHARPE, W. (1964) : “Capital Asset Prices : A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance*, 19(3), 425–442.
- SHEFRIN, H. (2005) : *A Behavioral Approach to Asset Pricing*. Elsevier Academic Press.
- SHEFRIN, H., ET M. STATMAN (2000) : “Behavioral Portfolio Theory,” *Journal of Financial and Quantitative Analysis*, 35(2), 127–151.

- SHILLER, R. (1981) : “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?,” *American Economic Review*, 71(3), 421–436.
- (1988) : “The Probability of Gross Violations of a Present Value Variance Inequality,” *Journal of Political Economy*, 96(5), 1089–92.
- (2000) : *Irrational Exuberance*. Princeton University Press.
- (2003) : “From Efficient Markets Theory to Behavioral Finance,” *Journal of Economic Perspectives*, 17(1), 83–104.
- (2005) : *Irrational Exuberance*. Princeton University Press, 2nd edn.
- SHIRYAEV, A. N. (1999) : *Essentials of Stochastic Finance*. World Scientific.
- SIMON, H. (1983) : *Reason in Humain Affairs*. Stanford University Press.
- SIMON, Y., ET P. JOFFRE (1997) : *Encyclopédie des marchés financiers* chap. 20, pp. 408–432. Economica, 2nd edn.
- SLOVIC, P. (1972) : “Psychological Study of Humain Judgment,” *Journal of Finance*, 27(4), 779–801.
- SLOVIC, P., ET W. BAUMAN (1972) : “Analyzing the Use of Information in Investment Decision Making : A methodological Proposal,” *Journal of Business*, 45(2), 283–301.
- SOLOMONOFF, R. (1960) : “A Preliminary Report on a General Theory of Inductive Inference,” Discussion paper, Cambridge, Report V-131.
- SORNETTE, D. (2002) : *Why Stock Market Crash : Critical Events in Complex Financial Systems*. Princeton University Press.
- SOSVILLA-RIVERO, S., F. FERNÁNDEZ-RODRIGUEZ, ET J. ANDRADA-FÉLIX (2005) : “Testing chaotic dynamics via Lyapunov exponents,” *Journal of Applied Econometrics*, 20(7), 911–930.
- STAMBAUGH, R. (1982) : “On the exclusion of assets from tests of the two-parameter model : A sensitivity analysis,” *Journal of Financial Economics*, 10(3), 237–268.
- STATTMAN, D. (1980) : “Book values and stock returns,” *The Chicago MBA : A Journal of Selected Papers*, 4, 25–45.
- SZAKMARY, A. C., ET D. B. KIEFER (2004) : “The Disappearing January/Turn of the Year Effect : Evidence From Stock Index Futures and Cash Markets,” *Journal of Futures Markets*, 24(8), 755–784.
- TAI, C.-S. (2003) : “Are Fama-French and momentum factors really priced?,” *Journal of Multinational Financial Management*, 13(4-5), 359–384.
- TAYLOR, A. M. R. (2003) : “Robust Stationarity Tests in Seasonal Time Series Processes,” *Journal of Business and Economic Statistics*, 21(1), 156–163.
- TAYLOR, S. J. (1994) : “Modeling Stochastic Volatility : A Review and Comparative Study,” *Mathematical Finance*, 4(2), 183–204.

- THALER, R. (1987) : “Anomalies : Seasonal Movements in Security Prices II : Weekend, Holiday, Turn of the Month, and Intraday Effets,” *Journal of Economic Perspectives*, 1(2), 169–177.
- (1999) : “The End of Behavioral Finance,” *Financial Analysts Journal*, 55(6), 12–17.
- TIROLE, J. (1982) : “On the Possibility of Speculation under Rational Expectations,” *Econometrica*, 50(5), 1163–81.
- TOBIN, J. (1958) : “Liquidity Preference as Behavior Toward Risk,” *Review of Economic Studies*, 25(2), 65–86.
- TSAY, R. S. (2002) : *Analysis of Financial Time Series*. John Wiley & Sons.
- (2005) : *Analysis of Financial Time Series*. Wiley-Interscience, 2nd edition edn.
- TURING, A. (1936) : “On Computable Numbers, with an Application to the Entscheidungsproblem,” *Proceedings of the London Mathematical Society*, 2(42), 230–65.
- TVERSKY, A., ET D. KAHNEMAN (1974) : “Judgment under Uncertainty : Heuristics and Biases,” *Science, New Series*, 185(4157), 1124–1131.
- VELUPILLAI, K. V. (2004) : “Economic Dynamics and Computation-Resurrecting the Icarus Tradition,” *Metroeconomica*, 55(2-3), 239–264.
- VERRECCHIA, R. E. (1975) : “A Proof of the Existence of "Consensus Beliefs",” *Journal of Finance*, 34(4), 957–963.
- VERYZHENKO, I., O. BRANDOUY, ET P. MATHIEU (2010) : “Agent’s Minimal Intelligence Calibration for Realistic Market Dynamics,” in *Progress in Artificial Economics*, ed. by M. Li Calzi, L. Milone, et P. Pellizzari, vol. 645 of *Lecture Notes in Economics and Mathematical Systems*, pp. 3–14. Springer Berlin Heidelberg.
- WALLIS, J., ET N. MATALAS (1970) : “Small Sample Properties of H and K-Estimators of the Hurst Coefficient h,” *Water Resources Research*, 6(6), 1583–1594.
- WALSH, K. (2006) : “Is the Ex Ante Risk Premium Always Positive? Further Evidence,” *Australian Journal of Management*, 31(1), 93–113.
- WASSERFALLEN, W., ET H. ZIMMERMAN (1985) : “The behavior of Intra-Daily Exchange Rates,” *Journal of Banking and Finance*, 9(1), 55–72.
- WERMERS, R. (2000) : “Mutual Fund Performance : An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses,” *Journal of Finance*, 55(4), 1655–1703.
- WERNDL, C. (2009) : “What Are the New Implications of Chaos for Unpredictability?,” *The British Journal for the Philosophy of Science*, 60(1), 195–220.
- WOLPER, P. (1991) : *Introduction à la Calculabilité*. InterEditions.

Bibliographie

- WORKING, H. (1934) : “Prices of Cash Wheat and Futures at Chicago Since 1883, Wheat Studies of the Food Research Institute,” *Stanford University*, november, 75–124.
- ZAROWIN, P. (1990) : “Size, Seasonality, and Stock Market Overreaction,” *Journal of Financial and Quantitative Analysis*, 25(1), 113–125.
- ZENIL, H., ET J. DELAHAYE (2011) : “An Algorithmic Information-Theoretic Approach to the Behaviour of Financial Markets,” *forthcoming in the Journal of Economic Surveys*, themed issue on “Nonlinearity, Complexity and Randomness”.
- ZHANG, G., B. EDDY PATUWO, ET M. Y. HU (1998) : “Forecasting with artificial neural networks : : The state of the art,” *International Journal of Forecasting*, 14(1), 35–62.
- ZHANG, G. P., ET V. L. BERARDI (2001) : “Time Series Forecasting with Neural Network Ensembles : An Application for Exchange Rate Prediction,” *Journal of the Operational Research Society*, 52(6), 652–664.
- ZVONKIN, A., ET L. LEVIN (1970) : “The Complexity of Finite Object and the Development of the Concepts of Information and Randomness by Mean of the Theories of Algorithms,” *Russ. Math Survey*, 25 (6), 83–124.